

Políticos españoles en la red social Twitter: actores principales, estrategias y sentimiento político en los medios de comunicación digital

Luis Ruiz Torres

Una tesis presentada en conformidad con los requerimientos
para el MSc en Economía, Finanzas y Computación

Universidad de Huelva y Universidad Internacional de Andalucía

uhu.es

un
i Universidad
Internacional
de Andalucía
A

Noviembre 2018

Políticos españoles en la red social Twitter: actores principales, estrategias y sentimiento político en los medios de comunicación digital

Luis Ruiz Torres
Máster en Economía, Finanzas y Computación

María Luisa Vílchez Lobato
Universidad de Huelva

2018

Abstract

The present work analyzes the behavior of different Spanish politicians in the social network Twitter. The most relevant actors and their influence are also examined. To do this, the network that forms their interactions is built and different global and local measures are studied. In the analysis we use the Gephi network visualization tool. In addition, the different strategies that political parties follow are indicated according to an analysis of communities in the network. This study also investigates the possible ideological closeness between some media and the different political parties. Through an analysis of the feelings of the news headlines, we approximate the objectivity of said means.

Resumen

El presente trabajo analiza el comportamiento de distintos políticos españoles en la red social Twitter. También se examina a los actores más relevantes y su influencia. Para ello, se construye la red que forman sus interacciones y se estudian distintas medidas globales y locales. En el análisis utilizamos la herramienta de visualización de redes Gephi. Además, se señalan las distintas estrategias que siguen los partidos políticos de acuerdo a un análisis de comunidades en la red. Este estudio también investiga la posible cercanía ideológica entre algunos medios de comunicación y los distintos partidos políticos. Mediante un análisis de sentimientos de los titulares referentes a noticias, aproximamos la objetividad de dichos medios.

Key words / Palabras clave: red, políticos, interacciones, comunidad, medios, centralidad, grafo.

Agradecimientos

Me gustaría agradecer a la profesora María Luisa Vílchez Lobato por su templanza y su esfuerzo para guiarme en el trabajo. También agradezco su paciencia y su cordialidad para atender en cualquier condición. Además, agradezco a Ester Andrés Villa por su constante apoyo. Finalmente, me gustaría agradecer a mi familia por su capacidad para no rendirse nunca ante la adversidad.

Tabla de Contenidos

1. Introducción	1.
2. Conceptos básicos sobre redes	3.
2.1 Definiciones básicas	3.
2.2 Medidas de análisis global	4.
2.3 Medidas de análisis local	7.
2.4 Modularidad	10.
3. Metodología	10.
4. Resultados	13.
4.1 Análisis global de la red en estudio	14.
4.2 Análisis local de la red en estudio	19.
4.3 Detección de comunidades en la red	29.
4.4 Análisis del sentimiento político en los medios de comunicación	34.
5. Conclusiones y posibles extensiones del trabajo	39.
6. Referencias	41.
Apéndice	43.

Lista de Tablas

Tabla 1. Grupo de estudio	11.
Tabla 2. Ranking de usuarios basado en el grado de entrada y salida	20.
Tabla 3. Grado de intermediación del grupo de estudio	22.
Tabla 4. Grado de cercanía del grupo de estudio	24.
Tabla 5. Ranking basado en la centralidad del vector propio	25.
Tabla 6. Ranking según la media de los retuits recibidos	28.
Tabla 7. Análisis de sentimiento de titulares de noticias de “ABC”	36.
Tabla 8. Análisis de sentimiento de titulares sobre noticias de “El Mundo”	36.
Tabla 9. Análisis de sentimiento de titulares sobre noticias de “El País”	37.
Tabla 10. Análisis de sentimiento de titulares sobre noticias de “eldiario.es”	37.
Tabla 11. Análisis de sentimiento de titulares sobre noticias de “Público”	38.

Lista de Figuras

Figura 1. Ejemplo de distancia entre nodos	6.
Figura 2. Red global del grupo de estudio	15.
Figura 3. Red de interacciones entre miembros del grupo de estudio	17.
Figura 4. Reciprocidad en la red de interacciones entre miembros del grupo de estudio	18.
Figura 5. Nivel de excentricidad del grupo de estudio	26.
Figura 6. Políticos con mayor nivel de centralidad	27.
Figura 7. Red de interacciones con medios de comunicación	30.
Figura 8. Red de interacciones entre miembros del grupo de estudio	32.
Figura 9. Red de interacciones de Percival Manglano con el resto de miembros del grupo de estudio	34.

1 Introducción

Con la irrupción de los nuevos partidos (Podemos y Ciudadanos) en el espectro político español a raíz de las elecciones europeas de 2014, el papel de la comunicación política en España se transformó. Los nuevos partidos hicieron de las redes su canal favorito para la transmisión de su mensaje político, además, la gran novedad de estos partidos respecto a los tradicionales recaía en la forma en la que basaron su comunicación dentro de las redes, una comunicación en ambas direcciones donde cualquier usuario podía participar (Chaves-Montero y otros, 2017). Este hecho ha llevado a los partidos tradicionales a redefinir sus estrategias dentro de este nuevo entorno de comunicación, al igual que los medios de comunicación tradicionales habían hecho previamente. Nunca la política había estado tan cerca de la sociedad. Este acercamiento entre políticos y ciudadanos ha tenido influencia a la hora de conseguir seguidores y votos durante las últimas elecciones. Para comprender los resultados de dichas elecciones tenemos que estudiar el comportamiento de los distintos partidos políticos en las redes sociales y observar de este modo los aspectos que han permitido que unos partidos ganaran influencia y otros la perdieran.

En este contexto, la red social Twitter ha jugado un papel fundamental. Esta red social nacida en 2006 y con más de 500 millones de usuarios, se caracteriza por la facilidad para establecer debates centrados en la actualidad social, política y deportiva. “Los estudios centrados en el debate político en *Twitter* y en sus efectos son sin duda los más abundantes en el campo de la comunicación política” (Campos-Domínguez, 2017: 787). Actualmente, España se encuentra en el top 10 de países donde más usuarios de Twitter existen, suponiendo el 2,6% del total. También destaca el hecho de que el 90% de los usuarios que acceden a Twitter lo hacen vía teléfono móvil. En esta red social los usuarios pueden seguir a otros y ser seguidos, además, distinguimos cuatro unidades distintas de interacción, el “tuit”, el “retuit”, la respuesta y el “me gusta”. El tuit consiste en un mensaje de 280 caracteres como máximo, éste lo podrán ver todos los usuarios que tengan la condición de seguidor, además de todos aquellos que accedan a tu perfil si éste es público. Por otro lado, el retuit consiste en compartir en tu perfil y para tus seguidores un mensaje de otro usuario. Generalmente esta interacción tiene carácter informativo o de apoyo. La respuesta es por definición un mensaje que contesta a otro. Y por último, el “me gusta” marca los tuits que te han gustado y los agrupa en una lista de “me gusta” en el perfil.

En el presente trabajo vamos a analizar el comportamiento de diversos políticos en la red social Twitter. Para ello, seleccionaremos hasta 45 políticos españoles que sirvan de representación de los partidos más importantes en el panorama nacional e incluiremos alguno, que no formando parte de alguno de éstos, pudiera cumplir con los criterios de selección. El objetivo de esta selección es obtener una representación del panorama político español en la red social.

En una primera parte del trabajo estudiaremos los aspectos globales y locales de la red que forman los políticos seleccionados en base a sus respuestas y retuits. Uno de los objetivos es identificar aquellos políticos más destacados en la red, los cuales interactúan más con el resto de usuarios y con el resto de políticos seleccionados. Además, queremos detectar cuáles son los usuarios a los que más referencia hace el conjunto de políticos, así como las fuentes de información que utilizan en la red. Por otro lado, también estamos interesados en conocer el comportamiento de los distintos partidos en la red, para observar cuáles de ellos son más cercanos a los usuarios y ver si su estrategia en la red es acertada.

En los últimos años hemos visto, tanto en las elecciones de Estados Unidos como en el Brexit, la proliferación de noticias falsas que afectaban a distintos candidatos en los momentos más álgidos de las campañas electorales, las cuales han marcado el rumbo de las propias elecciones. “Las informaciones falsas o trucadas (“fake news”), la propaganda política...suponen una amenaza para la democracia” (White, 2017: 8). La situación política en España tras el triunfo de la moción de censura al gobierno del Partido Popular augura un panorama donde podrían darse elecciones generales. Las redes sociales se han convertido en un medio de información más donde circulan noticias de todo tipo. En este contexto, los medios de comunicación tradicionales han evolucionado y su presencia digital tanto en la web como en las redes sociales ya es notoria. La influencia de los medios de comunicación en los votantes sigue siendo un hecho, si bien, ahora tienen la oportunidad de llegar a más ciudadanos a través de las redes sociales.

Nuestro objetivo para concluir el trabajo será observar el comportamiento de distintos medios, escorándolos primero en base al uso que los políticos hacen de ellos en la red social Twitter. Para ello, realizaremos un análisis de sentimiento con el fin de detectar la posible cercanía de algunos medios de comunicación con los distintos partidos. Además, analizaremos el sentimiento político general durante los primeros meses de gobierno del PSOE.

2 Conceptos básicos sobre redes

En este apartado vamos a exponer una serie de definiciones y medidas básicas sobre grafos y análisis de redes sociales, las cuales utilizaremos posteriormente para analizar nuestros resultados.

2.1 Definiciones básicas

Definición 2.1. Una *red social* es aquella red formada por un conjunto de actores (individuos, países, organizaciones...) los cuales están conectados por algún tipo de relación. Los actores se denominan *nodos* y las relaciones *enlaces*. En nuestro estudio los nodos se corresponden con el grupo de políticos seleccionados y el resto de usuarios a los que están conectados. Los enlaces se corresponden con las respuestas y retuits en la red social Twitter. Una forma de representar una red social es mediante un grafo, donde los nodos se corresponden con los vértices y los enlaces con las aristas del mismo.

Definición 2.2. En la teoría básica de grafos un *grafo* X se puede definir como “un conjunto de vértices, $V(X)$, y un conjunto de ejes, $E(X)$, donde cada eje es un par desordenado de vértices distintos de X ” (Godsil y Royle 2001: 1). Esta definición hace referencia a la definición de grafo no dirigido, el cual es considerado comúnmente como grafo en sentido general. Podemos definir los conjuntos del siguiente modo:

- $G = (V(G), A(G))$ donde G es el conjunto denominado Grafo y A es el conjunto de aristas.
- $V(G) \neq \emptyset$
- $A(G) \subseteq V(G) \times V(G)$

Definición 2.3. Los grafos que vamos analizar son *grafos dirigidos o digrafos*, en estos grafos el orden de los pares de vértices que forman las aristas tiene un significado especial y se representan con flechas que tienen un sentido o son bidireccionales con ambos vértices, por lo que una arista cualquiera (v, w) donde “ v ” es el vértice origen y “ w ” el vértice destino, es distinta de una arista (w, v) . Se trata de relaciones distintas de respuesta o retuit en nuestro caso, ya que no es lo mismo ser los autores o los receptores de dicha interacción.

Definición 2.4. En los grafos dirigidos distinguimos entre *grado de entrada* y *salida* de un vértice v . El grado de entrada lo podemos denotar como $g^-(v)$ y hace referencia al número de aristas cuyo sentido apunta a dicho vértice, es decir, el vértice v actúa como destino. Por otro lado, el grado de salida se refiere al número de aristas que parten del vértice v , actuando en este caso como vértice origen y lo denotamos como $g^+(v)$. La suma de ambos tipos da como resultado el grado de dicho vértice, por lo que $g^-(v) + g^+(v) = g(v)$. Además, podemos distinguir dos tipos de vértices dependiendo de su grado de entrada o salida. Si $g^+(v) = 0$ diremos que v es un vértice *sumidero* y si $g^-(v) = 0$ diremos que se trata de un vértice *fuentes*. Por último, si $g^-(v) = g^+(v) = 0$ denotamos al vértice v como *vértice aislado* (Rosen y otros, 2018).

Definición 2.5. El *orden de un grafo* se denota como $|V|$ y se corresponde con el número de vértices que posee.

Definición 2.6. Decimos que los enlaces o aristas tienen *peso* cuando dicha relación se produce más de una vez entre un par de nodos, es decir, hacemos referencia al tamaño de la relación. En nuestro estudio todos los enlaces tienen peso, ya que un mismo usuario puede responder o retuitear a otro varias veces, si bien, algunas métricas de análisis no lo toman en cuenta.

Definición 2.7. Un *camino* en un grafo se entiende como “una secuencia de vértices y aristas alternados que empiezan y terminan con un vértice...cada vértice incide a las dos aristas que le siguen y preceden en la secuencia” (Izquierdo y Hanneman, 2013: 11). Si bien, en un grafo dirigido los vértices que conectan las aristas en dicha secuencia son vértice origen con vértice destino. Por otro lado, cuando hablamos de un *camino geodésico* entendemos como tal el de menor longitud que conecta dos vértices, donde la longitud se corresponde con el número de aristas consecutivas entre el par de vértices que une el camino.

2.2 Medidas de análisis global

La *conectividad* es una de las características importantes que tenemos que tener en cuenta en nuestra red. Decimos que un grafo dirigido está conectado¹ si todos sus vértices pueden ser

¹ Chris Godsil y Gordon Royle se refieren a un digrafo conectado como débilmente conectado (2001: 29), si bien, generalmente se utiliza el término conectado para referirse a la definición que hemos aportado.

alcanzados desde cualquier vértice sin tomar en cuenta la dirección de las aristas, es decir, existe un *camino* entre cada par de vértices. Se trata de observar el dígrafo como un grafo normal y ver si todos sus vértices están conectados. (Godsil y Royle, 2001).

Cuando hablamos de dígrafos distinguimos entre dos tipos de conexiones:

- Conexión fuerte, ésta se produce cuando dos vértices se unen mediante las dos direcciones posibles. Dados los vértices v y w , estos estarán conectados de forma fuerte si existe un camino (v, w) y otro (w, v) . Un grafo estará fuertemente conectado si para el conjunto de pares de vértices que lo forman existen caminos en ambas direcciones.
- Conexión débil, en un grafo dirigido se entiende como tal la existencia de un único camino unidireccional entre dos vértices, es decir, dos vértices cualesquiera que se encuentran solamente unidos por una relación direccional, ya sea (v, w) o (w, v) para el caso de dichos vértices.

De las distinciones anteriores obtenemos el siguiente concepto. Decimos que un conjunto de vértices que se encuentran conectados entre sí de forma fuerte forman un componente fuerte (Ruohonen, 2013).

La *densidad* es una medida que nos ayuda a observar el grado de conectividad de nuestra red, ésta se corresponde con el total de aristas existentes en relación al total de aristas posibles (Wasserman y Faust, 1994). La expresamos del siguiente modo:

$$\Delta(G) = \frac{L(G)}{n(n-1)} \quad \text{donde } L \text{ es el número de aristas del grafo y } n \text{ el número total de vértices.}$$

Si se trata de un grafo dirigido el denominador es $n(n - 1)$ que coincide con el número total de aristas dirigidas posibles.

Otras medidas globales como el *diámetro* o la *distancia media de los caminos* nos ayudan a comprender el grado de cohesión y proximidad de la red en su conjunto. El *diámetro* de un grafo hace referencia a “la máxima distancia entre dos vértices distintos en el grafo” (Godsil y Royle,

2001: 16), donde la distancia se corresponde con la longitud del camino geodésico entre dichos vértices, veamos el siguiente ejemplo:

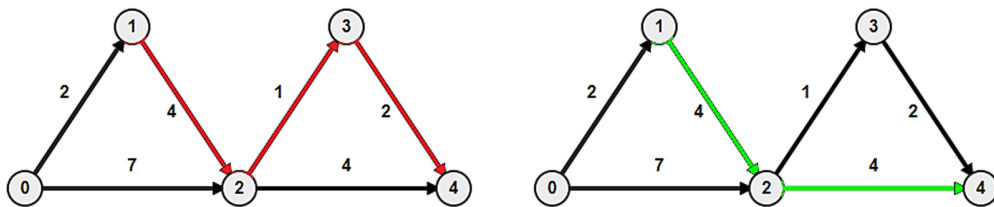


Figura 1. Ejemplo de distancia entre nodos. Elaboración propia con visualgo.net y Paint.

En la anterior figura observamos dos grafos dirigidos idénticos y con pesos, en ellos hemos señalado con distintos colores dos posibles longitudes entre los nodos “1” y “4”. En el grafo de la izquierda podemos ver en color rojo un camino de longitud 3 entre dichos vértices, en el de la derecha encontramos señalado en color verde un camino de longitud 2 que une los mismos vértices. La longitud del camino señalado en verde se corresponde con la distancia entre los vértices. En notación matemática podemos expresar el diámetro de un grafo “ G ” de la siguiente forma:

- $Diametro(G) = \max\{d(v, w) | v, w \in V(G)\}$ donde “ $d(v, w)$ ” es la distancia entre los vértices v y w .

La *distancia media de los caminos* se corresponde con la media de la distancia de todos los pares de vértices, es decir, la media de longitud de los caminos geodésicos que conectan cada par de nodos en la red (Barabási, 2014). Esta medida se expresa de la siguiente forma:

- $\bar{d}(G) = \frac{1}{n(n-1)} \sum_{v,w}^n d(v, w)$ donde $v \neq w$.

La última medida de análisis global de la red que vamos a exponer se corresponde con la *reciprocidad*. Ésta indica la proporción de conexiones bidireccionales sobre el total de las conexiones.

2.3 Medidas de análisis local

Las métricas relacionadas con la *centralidad* del nodo nos aproximan a la importancia de cada nodo dentro de la red. Las primeras nociones sobre la centralidad en el ámbito de las comunicaciones sociales aparecen con una investigación dirigida por el psicólogo Alexander Bavelas en el año 1948, a partir de entonces surgieron unos primeros estudios cuyas conclusiones ligaban el concepto de centralidad con el liderazgo, la resolución de problemas y la satisfacción (Freeman, 1978). Las medidas de centralidad que vamos a abordar se corresponden con: *Intermediación*, *Grado*, *Cercanía*, *Excentricidad* y *Centralidad del vector propio*. Los tres primeros conceptos fueron abordados por Freeman, el cual expone y detalla varias medidas basadas en dichos conceptos (1978). El primer concepto que vamos a exponer de nuestra red es el de la centralidad del grado, este aspecto ha sido desarrollado como una característica que ayuda a explicar la capacidad de comunicación que ejerce cada nodo en la red. “El grado de un punto se considera tan importante como un índice de su potencial actividad de comunicación” (Freeman, 1978: 221).

En las redes direccionales encontramos varios conceptos relativos a la implicación que puede tener el grado de un actor en la red. Encontramos que un usuario de la red cuyo grado de entrada es alto se denomina un usuario *prestigioso*, “el prestigio de un actor aumenta a medida que el actor se convierte en objeto de más lazos pero no, necesariamente, cuando es el mismo actor el que los inicia” (Wasserman y Faust, 1994: 196). Esta es una clara referencia al peso de las aristas, un nodo puede tener un grado de entrada con pesos muy alto pero relacionarse solo con uno o varios nodos de forma intensa lo cual no significa que dicho nodo sea prestigioso puesto que goza de pocos lazos. Por otro lado hallamos que un usuario con un grado de salida elevado se puede definir como un usuario *influyente*, si bien, tenemos que tener en cuenta el contexto en el que nos encontramos (Hanneman y Riddle, 2005).

La segunda medida de centralidad que vamos a exponer es la *intermediación*, esta medida trata de captar la centralidad midiendo el número de veces que un nodo se sitúa entre otro par de nodos, es decir, dado un usuario cualquiera en la red, observamos su nivel de intermediación como el porcentaje de veces que dicho usuario aparece en el medio del camino más corto entre otros dos usuarios. En el análisis de redes esta medida sirve para determinar la capacidad de control que tiene un nodo sobre la información que se transmite, puesto que un nodo con un

grado alto de intermediación se encuentra con frecuencia en el medio de distintas relaciones entre otros nodos, “betweenness es útil como un índice del potencial de un punto para el control de la comunicación” (Freeman, 1978: 224). En notación matemática se expresa de la siguiente forma:

$$- C_B(x) = \sum_{v,w \in V(G)}^n \frac{g_{v,w}(x)}{g_{v,w}} \quad \text{donde } v \neq w \neq x.$$

En la fórmula anterior tenemos varios elementos que vamos a detallar, el primero se corresponde con $g_{v,w}$, el cual hace referencia al total de caminos geodésicos que unen los nodos v y w , el segundo es $g_{v,w}(x)$ y se interpreta como el número de dichos caminos mínimos en los que el vértice x se sitúa entre ambos vértices. Por último, tenemos n como representación del número total de vértices en el grafo (Freeman, 1978). Esta medida se puede normalizar entre 0 y 1 atendiendo al número total de pares de vértices sin incluir el vértice x , siendo éste $(n - 1) \times (n - 2)$.

La tercera métrica de centralidad del nodo es la *cercanía*, esta medida consiste en la distancia a la que un nodo se encuentra del resto de nodos del grafo tomando en cuenta sus caminos geodésicos con dichos nodos. Se trata de observar si esa distancia es reducida para hablar de cercanía con el resto de nodos. Freeman repasa la obra de varios autores y expone dos puntos de vista sobre el papel que juega esta medida de centralidad, el primero de ellos es el de la independencia del nodo de acuerdo a la ausencia de intermediarios en sus caminos geodésicos con el resto de nodos, el segundo hace referencia a la eficiencia en términos de coste o tiempo en la comunicación (1978: 225). En notación matemática lo expresamos de la siguiente forma:

$$- C_C(v) = \frac{1}{\sum_{w \in V(G)}^n d(v,w)} \quad \text{donde } d(v,w) \text{ es la distancia geodésica y } v \neq w.$$

En la expresión anterior tenemos la inversa de la suma de las distancias mínimas entre el vértice v y el resto de vértices en el grafo. El concepto de distancia se corresponde con el explicado anteriormente en el análisis del diámetro de la red. También podemos definir la lejanía como la suma de las distancias geodésicas sin más que invertir la función anterior, quedándonos únicamente el denominador como expresión.

La *excentricidad* es otra medida de centralidad que está relacionada, al igual que la cercanía, con la distancia entre vértices. Esta medida es parecida al diámetro de la red pero aplicada a nivel individual entre usuarios. La podemos definir como la longitud del camino geodésico que separa al nodo más alejado de la red respecto al nodo que estamos analizando, es decir, si estamos analizando la excentricidad del nodo v estaremos calculando la distancia del nodo más lejano a él en la red (Izquierdo y Hanneman, 2006). En notación matemática lo denotamos del siguiente modo:

$$- C_E(v) = \max\{d(v, w) | w \in V(G)\}$$

Aquellos nodos que posean una mayor excentricidad serán aquellos más periféricos en la red. Se trata de una medida que puede servir de complemento a la centralidad basada en la cercanía de actores.

La última métrica de centralidad que vamos a exponer se corresponde con la *centralidad del vector propio* o *eigenvector centrality* en inglés. Esta medida es la primera que toma en cuenta la centralidad de los vecinos de un nodo para su cálculo. Intuitivamente decimos que un nodo será más central de acuerdo a la centralidad de sus vecinos. A diferencia del grado, esta medida valora a todos los nodos a los que se está conectado y por lo tanto va más allá de las relaciones meramente cuantitativas. La siguiente expresión se corresponde con su primera² versión:

$$- \lambda C_{VP}(v) = \sum_w R_{vw} C_{VP}(w) \text{ donde } R \text{ es la matriz de adyacencia}^3 \text{ y } \lambda \text{ es una constante.}$$

En la anterior fórmula observamos una medida de centralidad que pondera la matriz de adyacencia por la centralidad del resto de actores a los que estamos unidos. Conviene señalar que

² Phillip Bonacich en su trabajo *Power And Centrality: A Family of Measures* propone una segunda medida de vector propio basada en un nuevo parámetro denominado β . Este parámetro multiplica a C_{VP} y sirve para tener en cuenta otro tipo de relaciones donde una mayor centralidad de tus vecinos implica que tú seas menos central, teniendo en este caso signo negativo, o bien para dar más o menos valor a la centralidad de los vecinos (magnitud de β) (Bonacich, 1987).

³ La matriz de adyacencia se corresponde con una matriz cuadrada formada por las relaciones entre los vértices. Dada una matriz de un grafo dirigido D podemos denotar la matriz de adyacencia como R_D donde el elemento $R_D[v, w]$ se corresponde con el número de aristas que van desde el vértice v al w (Rosen y otros, 2000).

el vector C_{VP} está formado por el conjunto de centralidades de los vértices y es un autovector de la matriz de adyacencia R , además, la constante λ es su autovalor (Bonacich, 1987).

2.4 Modularidad

La *modularidad* consiste en la búsqueda de comunidades⁴ o módulos. Las comunidades son conjuntos de nodos con un alto grado de conectividad entre sí y nos ayudan a comprender el comportamiento de algunos grupos de usuarios en la red. En los grafos que vamos a presentar en el trabajo los nodos tendrán un color de acuerdo a la comunidad a la que pertenecen.

3 Metodología

Para la selección del grupo de estudio utilizado en el análisis de redes sociales en Twitter, nos hemos basado en el objetivo de conseguir una representación variada de los principales partidos políticos en España en la red social. Primero hemos asignado 17 puestos para los principales partidos de ideología conservadora (Partido Popular y Ciudadanos), y otros 17 para los principales partidos de ideología progresista (Partido Socialista y la confluencia Unidos Podemos-En Comú Podem- En Marea). Una vez hecho esto, se reparte la cifra impar de acuerdo a lo asignado en el grupo contrario, tomando como criterio la diferencia entre partidos nuevos y tradicionales.

Además, hemos asignado 3 puestos para cada uno de los partidos independentistas catalanes con representación en el congreso (Partido Demócrata Europeo Catalán (Grupo mixto) y Esquerra Republicana de Catalunya), y otros dos puestos para el Partido Nacionalista Vasco. Por último, hemos incluido un puesto para EH Bildu con dos escaños en el congreso.

⁴ En la búsqueda de comunidades Gephi utiliza un algoritmo que establece una comunidad por cada nodo en la red y después mide si aumenta el nivel de modularidad cuando metemos al nodo original en la comunidad del vecino. El proceso se repite secuencialmente hasta que el cambio de comunidad pueda conseguir una mejora en la modularidad. Para medir la modularidad utilizan una formula basada en la densidad de las conexiones dentro de cada comunidad en comparación con las conexiones inter-comunidades. Una vez que se alcanza la máxima modularidad se establece una nueva red donde las comunidades obtenidas son ahora nodos y el peso de las aristas entre dichos nodos (comunidades) se corresponde con la suma del peso de las aristas que unen nodos inter-comunidades, seguidamente se vuelve a aplicar la primera parte (Blondel y otros, 2008).

Tabla 1. Grupo de estudio.

<i>Nombre</i>	<i>Partido</i>	<i>Nombre</i>	<i>Partido</i>
<i>Mariano Rajoy</i>	PP	<i>Manuela Carmena</i>	Podemos
<i>Soraya Sáenz de Santamaría</i>	PP	<i>Ada Colau</i>	Podemos
<i>María Dolores de Cospedal</i>	PP	<i>Juan Carlos Monedero</i>	Podemos
<i>Fernando Martínez Maíllo</i>	PP	<i>Ramón Espinar</i>	Podemos
<i>Pablo Casado Blanco</i>	PP	<i>Albert Rivera</i>	Ciudadanos
<i>Rafael Hernando</i>	PP	<i>Inés Arrimadas</i>	Ciudadanos
<i>Xavier Albiol</i>	PP	<i>Félix Álvarez</i>	Ciudadanos
<i>Javier Maroto</i>	PP	<i>Juan Carlos Girauta</i>	Ciudadanos
<i>Percival Manglano</i>	PP	<i>Toni Cantó</i>	Ciudadanos
<i>Pedro Sánchez</i>	PSOE	<i>José Manuel Villegas</i>	Ciudadanos
<i>Susana Díaz</i>	PSOE	<i>Fran Hervías</i>	Ciudadanos
<i>Miquel Iceta</i>	PSOE	<i>Fernando de Páramo</i>	Ciudadanos
<i>Patxi López</i>	PSOE	<i>Gabriel Rufián</i>	ERC
<i>José Luis Ábalos</i>	PSOE	<i>Oriol Junqueras</i>	ERC
<i>Ximo Puig</i>	PSOE	<i>Marta Rovira</i>	ERC
<i>Emiliano García Page</i>	PSOE	<i>Carles Puigdemont</i>	PdCAT
<i>Adriana Lastra</i>	PSOE	<i>Quim Torra</i>	PdCAT
<i>Pablo Iglesias</i>	Podemos	<i>Jordi Turull i Negre</i>	PdCAT
<i>Irene Montero</i>	Podemos	<i>Íñigo Urkullu</i>	PNV
<i>Íñigo Errejón</i>	Podemos	<i>Aitor Esteban</i>	PNV
<i>Pablo Echenique</i>	Podemos	<i>Arnaldo Otegi</i>	EH Bildu
<i>Alberto Garzón</i>	Podemos	<i>Santiago Abascal</i>	VOX
		<i>Rosa Díez</i>	-

Una vez distribuidos los puestos, hemos utilizado dos criterios complementarios y no excluyentes para seleccionar a los políticos, los cuales se corresponden con la popularidad en la red social en base a un mínimo de seguidores (20.000) y la importancia dentro del partido en base al cargo que ocupan en el partido (actual o pasado reciente). Adicionalmente, se han añadido dos políticos de partidos que no guardan representación en el congreso, pero que, por su elevada popularidad y debido al papel que jugaron o podrían jugar en la política española futura, se han sumado al grupo de estudio (Rosa Díez y Santiago Abascal).

Para la extracción y depuración de los tuits de los miembros del grupo de estudio, hemos utilizado el software libre R mediante el entorno de desarrollo integrado Rstudio. Este software de carácter estadístico nos permite conectarnos a la API de Twitter a través de la librería “twitterR”. Para conectarnos a dicha API, necesitamos de una clave API y unos códigos de acceso

que se obtienen registrando una cuenta en la plataforma de Twitter para desarrolladores. Una vez que obtenemos las claves, ya solo nos queda conectarnos a la API y cargar una lista con el nombre de usuario de los miembros del grupo de estudio. Después utilizamos la función “userTimeline” para extraer 3000 interacciones por usuario. Cuando ya tenemos todo los datos en el mismo panel, procedemos a su depuración y mantenemos únicamente las respuestas y retuits. Por último, formateamos ambos para acondicionarlos a la herramienta de visualización de redes Gephi.

Con el fin de analizar la red formada con las respuestas y retuits, procedemos a cargar los mismos en Gephi. Una vez cargados, comenzamos el diseño de la red utilizando uno de los algoritmos de distribución para dibujar el grafo que la representa (Force Atlas 2). Este algoritmo es el que mejor adecúa el aspecto del grafo para su correcta visualización. Teniendo el grafo dibujado, procedemos a ejecutar una serie de métricas que aparecen en el apartado de “Estadísticas”. Estas métricas nos ayudan a diseñar algunos aspectos del grafo, como el tamaño de los nodos o el color de los mismos, de acuerdo a las exigencias. Además, dichas métricas también nos sirven para realizar el análisis global y local de nuestra red. Por último, mediante la opción “Filtro” que incorpora, podemos diseñar diversas filtraciones de acuerdo a varios operadores lógicos y conseguir de este modo destacar algunos aspectos interesantes que comentaremos en los resultados.

En la última parte del trabajo hemos realizado un análisis de sentimientos de los titulares de noticias correspondientes a cinco medios de comunicación distintos. Los medios han sido seleccionados atendiendo a la cercanía en la red con una ideología progresista o conservadora. Si el medio es usado mayoritariamente por políticos de un espectro ideológico concreto, diremos a priori, que el medio pertenece a esa ideología. Es decir, seleccionamos el medio según la ideología de los políticos que más lo usan como fuente de información en la red social Twitter. El objetivo es obtener una muestra representativa de las principales ideologías políticas imperantes. Para observar cuáles son los medios que se sitúan más en un lado o en otro, filtramos con Gephi y buscamos dos medios que sean usados mayoritariamente por miembros de partidos progresistas y otros dos por conservadores. Por último, buscamos un medio que se sitúe entre ambas ideologías. Los medios seleccionados finalmente son “eldiario.es”, “Público”, “El País”, “ABC” y “El Mundo”. Los dos primeros son usados principalmente por políticos de ideología progresista, los dos últimos por el contrario son utilizados por políticos conservadores y “El

País” aparece como el medio que representa el centro ideológico, siendo utilizado por miembros de ambos espectros, si bien, destaca el uso que hacen de dicho medio los miembros de Ciudadanos.

Para llevar a cabo la última parte del trabajo hemos utilizado el lenguaje de programación Python a través del entorno de desarrollo integrado conocido como IDLE. Lo primero que hemos realizado son los códigos correspondientes con la extracción de los datos. Mediante la técnica conocida como Web Scraping hemos creado un código para extraer los titulares de la web de cada uno de los medios. A lo largo de tres meses hemos ido recopilando titulares de ámbito político. El segundo paso ha consistido en la creación de un código personalizado para clasificar los titulares de acuerdo a uno de los cuatro grandes partidos políticos del panorama nacional (PP, PSOE, Podemos y Ciudadanos). En dicho código hemos creado un diccionario con palabras que pueden representar a distintos partidos y hemos planteado la solución a diversos problemas que nos iban surgiendo, como la aparición de titulares en donde se hace referencia a distintos partidos o titulares que se refieren a gobiernos de otros países. Además, los distintos titulares se han clasificado atendiendo la primera aparición en cada titular de una palabra correspondiente al diccionario y se han tenido en cuenta aquellos titulares que se refieren a frases que transmiten los partidos o sus miembros. Por último, para realizar el análisis de sentimiento hemos elaborado un código para utilizar la API de análisis de sentimiento de Meaning Cloud (servicio de software de análisis semántico), la cual nos permite un máximo de 20000 solicitudes de análisis con la licencia gratuita.

Todos los códigos correspondientes a la extracción y depuración de datos en R, así como los pertenecientes a la extracción y análisis de sentimiento en Python, se encuentran detallados en el Apéndice.

4 Resultados

En este apartado vamos a analizar los resultados obtenidos, para ello expondremos algunas interpretaciones y consecuencias de los mismos. Además, tras su presentación gráfica incluiremos si fuera necesario algunas aclaraciones sobre su forma.

Los grafos que vamos a presentar se corresponden fundamentalmente con dos análisis diferenciados pero complementarios, el primero de ellos se corresponde con la red global de

interacciones del grupo de estudio y el segundo con la red de interacciones exclusivamente entre los miembros del grupo de estudio. Por último, en el apartado de detección de comunidades incluiremos algunos grafos más para resaltar algunas características en concreto, si bien, siempre serán grafos realizados a partir de los dos primeros. Los datos utilizados son siempre los mismos y se han detallado en el apartado de metodología. Cabe resaltar, que todas las métricas que vamos a analizar se corresponden con la red que representa la Figura 2 (red de interacciones globales mostrada en la siguiente página), salvo que se exprese lo contrario.

4.1 Análisis global de la red en estudio

En este primer análisis de la red nos centraremos en su estructura global, dejando para el siguiente apartado las características individuales de los actores de la red.

Todas las relaciones en nuestra red tienen peso y sentido (es una red dirigida), si bien, no todas las medidas que vamos a analizar tienen en cuenta dichas características. Además, estas propiedades no son necesarias en todos los aspectos de nuestro análisis.

El primer grafo que vamos a presentar se corresponde con el conjunto global de las respuestas y retuits que realizan los políticos que forman parte del grupo de estudio seleccionado. Cuando decimos global nos referimos al hecho de que el destino de esas interacciones puede ser cualquier usuario registrado en la red social Twitter. Más adelante, estudiaremos las interacciones que tienen lugar únicamente entre miembros del grupo de estudio.

Lo primero que nos encontramos tras cargar los datos en la herramienta Gephi, es el número de *nodos* que forman la red aparte de los 45 correspondientes al grupo de estudio, así como el número de relaciones distintas entre los nodos (número de aristas). La red se encuentra formada por un total de 13261 nodos y 20701 aristas por lo que el *orden del grafo* es de 13261. En nuestro caso solo los nodos pertenecientes al grupo de estudio tienen grado de entrada y salida, el resto de nodos únicamente poseen grado de entrada (nodos sumidero).

Notamos que el grafo ha sido diseñado atendiendo a ciertas características de la red que conviene detallar, el tamaño de los nodos se corresponde con el grado con *peso* de los mismos, normalizado en una escala 0.5-100 (suma del grado de entrada y salida con pesos). El peso podemos hacerlo notar en el tamaño del nodo y destacar de este modo aquellos usuarios que tienen un mayor número de interacciones con el resto. Debemos tener en cuenta que, si bien el tamaño de los nodos está diseñado en base al grado total con pesos, esto no quiere decir que dichos nodos tengan un comportamiento dinámico en la red, es decir, podemos encontrarnos nodos cuyo número de interacciones sea muy elevado pero dichas interacciones sean con un

grupo pequeño de individuos, produciéndose en este caso una concentración de la actividad del nodo en la red. Además, el color de los nodos se corresponde con la comunidad a la que dicho nodo pertenece de acuerdo a la propiedad de *modularidad* y el color de las aristas viene determinado por el color del nodo origen de la interacción. Por último, las etiquetas que aparecen se corresponden con aquellos usuarios que tienen un grado de entrada con pesos mayor, entre los que aparecen nuevos nodos en adicción a los pertenecientes al grupo de estudio, además, se han eliminado las etiquetas de los nodos correspondientes a medios de comunicación o periodistas con los que interactúan los políticos seleccionados, con el objetivo de aclarar el grafo y observarlos mejor en otra figura posteriormente.

Si analizamos la *conectividad* podemos afirmar que la red que tenemos en la Figura 2 es una red débilmente conectada o conectada, ya que existen caminos que unen cada par de nodos de la red sin tomar en cuenta el sentido. En su conjunto esta red forma un componente débil y gigante (componente más grande conectado en la red), además existe un componente⁵ conectado fuertemente, el cual está formado por el conjunto de políticos pertenecientes al grupo de estudio con la exclusión de Arnaldo Otegui y Emiliano García-Page. Esto demuestra que hemos escogido el grupo de estudio de forma adecuada y podemos obtener algunas conclusiones interesantes de cara a esta red.

Fijándonos en la red que estamos analizando podemos ver que la mayoría de los nodos solo tienen una conexión directa de entrada, por lo que no tiene sentido medir la *densidad*, pues el número de aristas posibles está muy alejado de las que existen realmente. Sin embargo, si analizamos esta medida tomando en cuenta la red que forman las interacciones de los miembros del grupo de estudio entre ellos mismos (la cual mostramos a continuación como Figura 3 y no es más que una submuestra de la red global (Figura 2)), obtenemos un nivel de densidad del 0'17. En dicha red solo se dan el 17% de las relaciones posibles, esto es debido en gran parte a la tipología social de la red (red de respuestas y retuits), además, es normal que las relaciones

⁵ Gephi interpreta como componente fuerte a cada nodo que posee únicamente grado de entrada o únicamente grado de salida, por ello, los resultados ofrecen más de doce mil componentes fuertes, sin embargo, un nodo que solo posee grado de entrada o grado de salida no puede alcanzar y ser alcanzado por el mismo nodo. Cabe destacar que se interpreta como componente conectado fuertemente a aquel más grande posible, por lo que los consiguientes componentes fuertes que puedan existir dentro de un mismo componente fuerte no son tomados en cuenta.

tengan lugar mayormente entre miembros de un mismo partido político como veremos en el análisis de la *modularidad*.

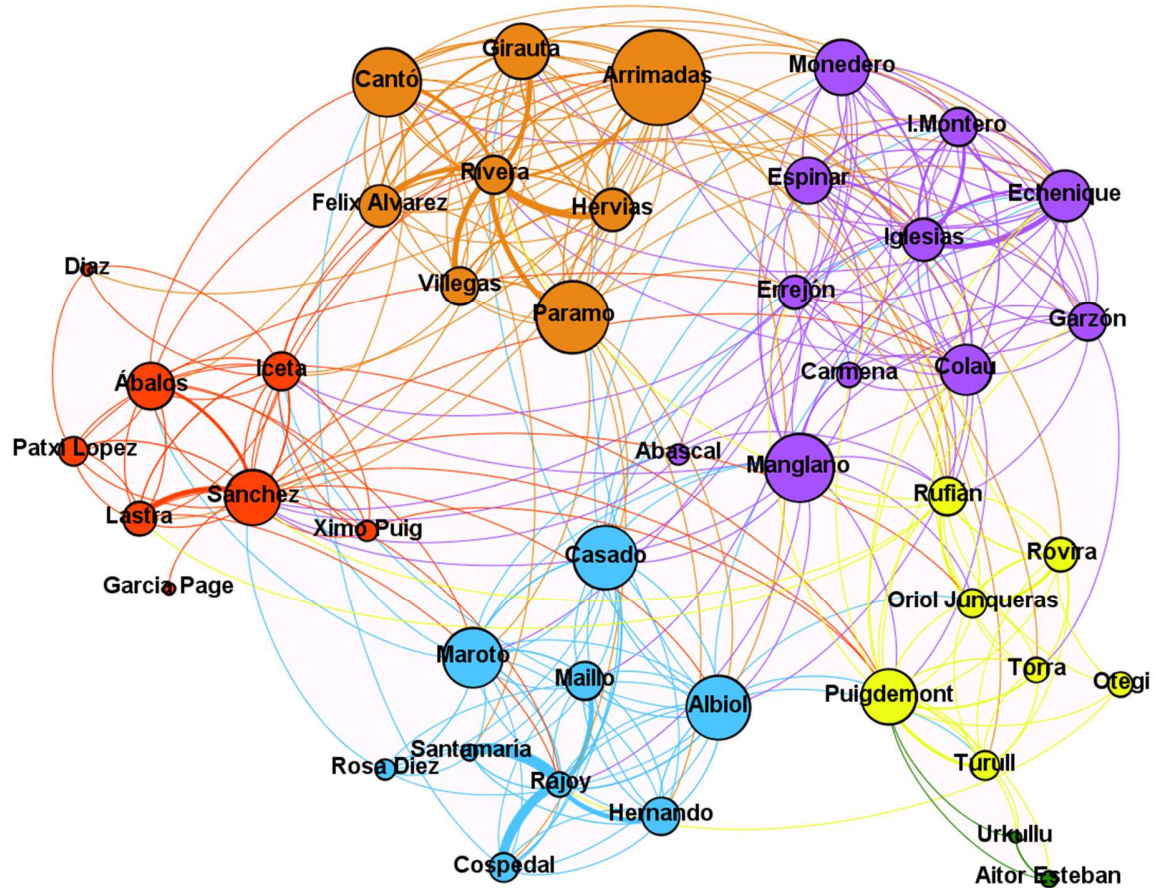


Figura 3. Red de interacciones entre miembros del grupo de estudio. Fuente: Elaboración propia con Gephi.

En la figura anterior, el tamaño de los nodos se corresponde con el grado de salida y el color de los mismos con la propiedad de modularidad. Además, el color de los enlaces se corresponde con el color del nodo origen de la interacción y su grosor con el peso del enlace.

Continuando con la red de la Figura 2, tenemos que el diámetro es de 6, es decir, son seis las aristas que separan al par de nodos más alejado de la red. La media de los caminos que existen en el grafo es de 3'1, por lo que son tres los enlaces que forman en media los caminos geodésicos de cada par de usuarios en la red. Dadas estas dos últimas medidas podemos señalar que la proximidad del conjunto de los nodos en la red es elevada. Esto se debe a los políticos que interactúan más con miembros de partidos distintos al que pertenecen. Cabe destacar que para el

cálculo de esta medida Gephi no toma en cuenta el conjunto de los caminos posibles sino solo los caminos geodésicos que existen en el grafo, por lo que si dos nodos no se encuentran conectados, esa posible conexión que podría existir no es tomada en cuenta en el denominador.

A continuación, vamos a observar la reciprocidad en los enlaces entre los políticos del grupo de estudio, ya que no podemos observar esta propiedad con el resto de nodos (sumidero) debido a la estructura de los datos extraídos.

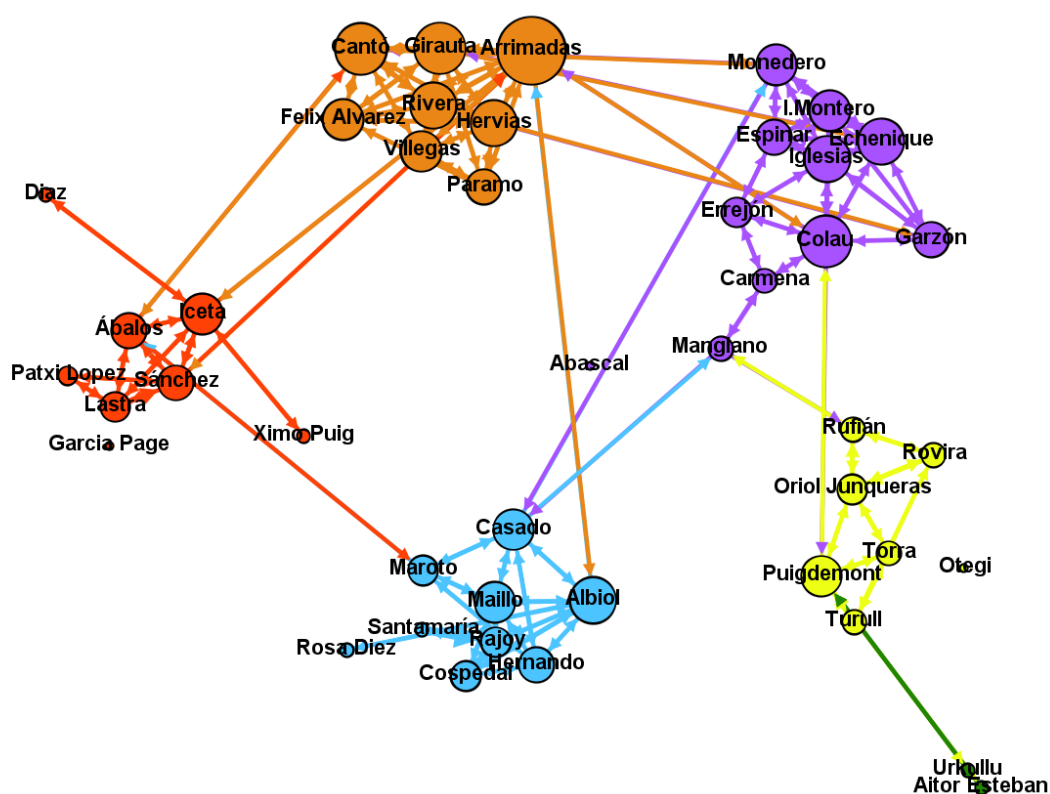


Figura 4. Reciprocidad en la red de interacciones entre miembros del grupo de estudio. Fuente: Elaboración propia con Gephi.

En la figura anterior solo aparecen las conexiones recíprocas. El resto de aristas se han eliminado, además, el tamaño de los nodos se corresponde con el grado y el color con la *modularidad*.

El 60% de las conexiones de la red son recíprocas y éstas se producen en mayor medida entre miembros del mismo partido, prácticamente todos los miembros de un mismo partido tienen conexiones bidireccionales con el resto de miembros. Cabe destacar que el mayor número de

interacciones recíprocas con miembros de partidos políticos distintos tienen lugar en Ciudadanos y Podemos con 7 y 8 respectivamente. Este aspecto es importante ya que una comunicación bidireccional puede implicar una mayor calidad de las interacciones, la cual ayuda a “consolidar ciudadanos en seguidores” (Chaves-Montero y otros, 2017: 20).

En el siguiente apartado vamos a continuar explicando propiedades de la red (Figura 2), si bien, esta vez nos centraremos en aspectos relacionados con los nodos.

4.2 Análisis local de la red en estudio

En la red que muestra la Figura 2 aparecen etiquetados nuevos nodos en adicción a los del grupo de estudio, éstos pertenecen a los principales partidos políticos y algunas instituciones públicas que poseen una cuenta en Twitter, como es el caso de los Ayuntamientos de Madrid o Barcelona, además de varias cuentas asociadas al Gobierno Central, como el Ministerio de Defensa o el Congreso de los Diputados. Estos nodos suelen recibir muchas interacciones con los miembros de los partidos que gobiernan dichas instituciones, y a pesar de que no podemos observar su grado de salida, aparecen entre los nodos con un grado total con peso mayor.

Debido a que tratamos con retuits y respuestas, gran parte de las interacciones son una muestra de apoyo o coincidencia en el punto de vista con otro usuario (retuit), estamos destacando lo que otro usuario dice, estableciendo así una relación donde la influencia tiene lugar en sentido inverso. En nuestra opinión, debemos interpretar el grado de entrada como medida de prestigio o importancia, añadiendo también cierto carácter influyente debido al contexto de nuestra red. De otro modo, el grado de salida debe observarse como una muestra de la diversidad o concentración de la actividad del político con el resto de usuarios⁶, así como de la cercanía ideológica entre políticos.

A continuación, detallamos los resultados sobre el grado de entrada y salida en nuestra red.

⁶ Un mayor número de enlaces entre los ciudadanos y los políticos ayuda al “acercamiento” entre ambos y reduce la “desafección política” (Chaves-Montero y otros, 2017: 21). De acuerdo a este supuesto, debemos interpretar un mayor grado de salida como una característica deseable para atraer seguidores y por lo tanto posibles votantes. Sin embargo, debemos tomar en cuenta que al interactuar demasiado con otros usuarios corremos el peligro de conseguir el efecto contrario y podríamos atraer a usuarios muy críticos además de otros denominados “troll”, los cuales se van a dedicar continuamente a realizar críticas y burlas reduciendo en muchos casos la calidad de nuestras interacciones.

Tabla 2. Ranking de usuarios basado en el grado de entrada y salida.

<i>Usuario</i>	<i>Grado de entrada</i>	<i>Usuario</i>	<i>Grado de salida</i>
<i>Europapress</i>	33	<i>Gabriel Rufián</i>	1061
<i>El País</i>	29	<i>Javier Maroto</i>	970
<i>eldiario.es</i>	26	<i>Aitor Esteban</i>	868
<i>Al Rojo Vivo</i>	26	<i>Quim Torra</i>	826
<i>Tve 24h</i>	25	<i>Carles Puigdemont</i>	696
<i>Espejo Público</i>	23	<i>José Luis Ábalos</i>	694
<i>El Periódico</i>	22	<i>Percival Manglano</i>	676
<i>La SER</i>	22	<i>José Manuel Villegas</i>	618
<i>Público</i>	22	<i>Íñigo Errejón</i>	603
<i>El Mundo</i>	21	<i>Adriana Lastra</i>	603
<i>Onda Cero</i>	21	<i>Rosa Díez</i>	256
<i>La Vanguardia</i>	20	<i>Miquel Iceta</i>	247
<i>sexta Noticias</i>	20	<i>Pedro Sánchez</i>	234
<i>Guardia Civil</i>	20	<i>Susana Díaz</i>	187
<i>Congreso</i>	19	<i>F.M. Maíllo</i>	182
<i>Ana Pastor</i>	19	<i>S.S. Santamaría</i>	174
<i>Ignacio Escolar</i>	19	<i>M.D. Cospedal</i>	152
<i>MasDeUno</i>	18	<i>Mariano Rajoy</i>	151
<i>ElHuffPost</i>	18	<i>Íñigo Urkullu</i>	123
<i>El Español</i>	17	<i>Rafael Hernando</i>	95

En la figura anterior encontramos dos tablas con un ranking sobre el grado de los nodos, en primer lugar tenemos aquellos veinte usuarios que poseen un grado de entrada mayor, es necesario recordar que el grado de entrada máximo que podemos encontrar en este análisis es de 45, ya que es el número de usuarios pertenecientes al grupo de estudio y son ellos los únicos de los que se posee información de salida (respuestas y retuits), por ello, cualquier usuario de la red solo puede ser retuiteado o respondido por 45 individuos distintos como límite. En la tabla de la derecha observamos veinte políticos entre los que separamos los diez con mayor grado de salida y los diez con menor grado de salida en la red, quedando en medio el resto de políticos pertenecientes al estudio que no aparecen en la tabla. Mostrar el grado completo no tiene sentido ya que obtendríamos el mismo ranking obtenido con el grado de salida, puesto que la suma de ambos grados (salida y entrada) arroja resultados similares, dado el rango en el que se sitúa el grado de entrada (0-45).

Los resultados sobre el grado de entrada nos aportan una clara conclusión, la importancia de los medios de comunicación en la red formada a partir de las interacciones de los políticos españoles es notoria. Salvo la Guardia Civil, el ranking de los veinte usuarios con mayor grado de entrada queda compuesto por medios de comunicación y contenido de los mismos, así como algunos periodistas. Destaca por encima de todos la agencia de noticias española Europapress, siendo alguna vez retuiteada o respondida por el 73% de los políticos pertenecientes al grupo de estudio, le siguen medios como El País y “eldiario.es”, habiendo tenido interacciones de entrada con 29 y 26 políticos distintos respectivamente. En definitiva, podríamos señalar que una noticia emitida en uno de estos medios llegaría antes al conjunto de los políticos que otra emitida en Ok Diario o El Confidencial.

El grado de entrada con pesos también puede servir para analizar la importancia de los medios, sin embargo, tal y como ya explicamos, son algunas instituciones y partidos políticos los principales nodos concentradores de interacciones de entrada. Hemos analizado el grado sin pesos con el objetivo de encontrar nuevas conclusiones sobre usuarios prestigiosos o prominentes, ya que la concentración del peso en aristas que se dirigen a partidos políticos u otros políticos que pertenecen al mismo partido se analizará cuando estudiemos la modularidad. Además, esta medida no nos provee de grandes conclusiones sobre la centralidad debido a que todas las medidas que nos aporta Gephi han sido diseñadas sin valorar el peso de las relaciones, y como señalamos anteriormente, el grado con pesos puede fallar a la hora de medir el prestigio de los usuarios en la red.

Si observamos el grado de salida concluimos que entre los políticos con mayores interacciones de salida tenemos una representación equitativa de los principales partidos políticos. Por otro lado, hay una clara dominación del Partido Popular y el Partido Socialista por la parte baja de la tabla. Entre todos los políticos del grupo de estudio destaca Gabriel Rufián cuya actividad en las redes sociales es generalmente conocida y aparece como el político que más se relaciona con usuarios distintos, con un total de 1061. Estas cifras son bastante significativas si tenemos en cuenta que solo disponemos de 3000 tuits por político y no todos son respuestas o retuits.

El grado es una medida sencilla que puede aproximarnos cómo de central es un nodo en la red, si bien, existen algunas debilidades que son corregidas por otras medidas. El principal problema del grado como medida de centralidad es la localidad del análisis de las relaciones, ya que es una medida que no tiene en cuenta la importancia de las relaciones que tienen sus vecinos y se centra

en relaciones directas. En nuestro caso la pregunta que nos deberíamos hacer es si consideramos más importante que un político tenga muchas interacciones en la red o que las interacciones sean con usuarios centrales en la red, cantidad o calidad. Nosotros consideramos que un equilibrio entre ambas es lo adecuado. Además, debemos distinguir la calidad de la interacción de la calidad o importancia del usuario con el que interactuamos.

Ahora vamos a exponer los resultados de intermediación para el grupo de estudio, puesto que el conjunto de nodos que conforma la red son, salvo el grupo de estudio, nodos sumidero.

Tabla 3. Grado de intermediación del grupo de estudio.

<i>Usuario</i>	<i>Intermediación</i>	<i>Usuario</i>	<i>Intermediación</i>
<i>C. Puigdemont</i>	0,000588	<i>Oriol Junqueras</i>	0,000128
<i>X.G. Albiol</i>	0,000423	<i>Pablo Casado</i>	0,000126
<i>Gabriel Rufián</i>	0,000401	<i>Jordi Turull</i>	0,000116
<i>Pedro Sánchez</i>	0,000387	<i>Ximo Puig</i>	0,000116
<i>I. Arrimadas</i>	0,000361	<i>F. Páramo</i>	0,000104
<i>Ada Colau</i>	0,000308	<i>J.M. Villegas</i>	0,000096
<i>Javier Maroto</i>	0,000275	<i>Marta Rovira</i>	0,00009
<i>Miquel Iceta</i>	0,000239	<i>Patxi López</i>	0,00008
<i>J.L. Ábalos</i>	0,000222	<i>Rosa Díez</i>	0,000077
<i>P. Echenique</i>	0,000206	<i>Mariano Rajoy</i>	0,000075
<i>P. Manglano</i>	0,000202	<i>F.M. Maíllo</i>	0,000073
<i>Toni Cantó</i>	0,000192	<i>Fran Hervías</i>	0,000057
<i>J.C. Monedero</i>	0,000177	<i>Ramón Espinar</i>	0,000056
<i>J.C. Girauta</i>	0,000171	<i>Félix Álvarez</i>	0,00005
<i>Íñigo Errejón</i>	0,000171	<i>M.D. Cospedal</i>	0,000044
<i>Aitor Esteban</i>	0,000171	<i>S. Abascal</i>	0,000042
<i>Quim Torra</i>	0,000162	<i>Irene Montero</i>	0,000041
<i>A. Garzón</i>	0,000152	<i>Susana Díaz</i>	0,000036
<i>A. Lastra</i>	0,000147	<i>Íñigo Urkullu</i>	0,000022
<i>Albert Rivera</i>	0,000145	<i>S.S Santamaría</i>	0,000021
<i>M. Carmena</i>	0,000129	<i>R. Hernando</i>	0,000013
<i>Pablo Iglesias</i>	0,000128	<i>G. Page/Otegui</i>	0

La tabla anterior nos ofrece el grado de intermediación normalizado de los políticos analizados. Los resultados obtenidos son muy bajos, sin embargo, esto no influye en nuestro análisis donde lo que importa es la posición que ocupan. Es sencillo comprender que una red formada en su mayoría por nodos sumidero, en la cual el 77,7% del total de los nodos solo posee 1 grado de

entrada, arrojará resultados de intermediación bajos debido al gran número de pares de nodos que se encuentran relacionados directamente sin ningún intermediario. Esto no afecta al análisis de la intermediación, ya que de eliminar dichos nodos los resultados serían muy parecidos.

A través de la intermediación detectamos la importancia de los políticos catalanes en la red, entre los primeros seis políticos con mayor intermediación cinco son catalanes. Destaca Carles Puigdemont como el político más central de la red, el cual se sitúa como elemento intermediador entre distintas conexiones en la red. Por la parte baja del ranking encontramos los políticos más aislados, los cuales no gozan de ninguna posición de control, como son el socialista Emiliano García-Page, Arnaldo Otegui, Rafael Hernando o Soraya Sáez de Santamaría entre otros.

Debemos tener presente el diseño de la red, pues se trata de una red de interacciones de los políticos españoles, por lo que la comunicación entre dos nodos no depende del intermediador, dicha comunicación tendrá lugar o no tendrá lugar, pero siempre será directa. Podemos entonces interpretar el papel del intermediador como un puente entre comunidades distintas, una especie de representante ideológico que actúa de unión entre ideologías parecidas en regiones distintas. De acuerdo con los últimos acontecimientos vividos entre finales del año 2017 y la actualidad, el panorama político catalán se ha podido ver desplazado a una posición privilegiada en la actualidad social, por lo que es posible que sus principales representantes hayan visto incrementada su centralidad en la red debido a un factible incremento de sus interacciones de entrada, aumentando así su nivel de intermediación al situarse en un mayor número de caminos geodésicos.

Otra de las métricas para la que hemos obtenido resultados es la cercanía. Antes de mostrar la tabla con los resultados, debemos aclarar que en las redes dirigidas se pueden calcular dos tipos de cercanía: de entrada y de salida. Esta última es la que nos ofrece Gephi. De acuerdo a esta aclaración, los índices de cercanía que vamos a mostrar se calculan tomando en cuenta únicamente los caminos de salida, es decir, medimos cómo de cerca se encuentra cada político del resto de actores de la red partiendo desde él. Esta medida se puede interpretar en nuestro contexto como una forma de medir la calidad de las relaciones directas de un actor. Un político que interactúa con otros políticos en la red puede crear debates interesantes que atraen al resto de usuarios, generando así posibles seguidores. A continuación, mostramos los resultados para esta métrica.

Tabla 4. Grado de cercanía del grupo de estudio.

<i>Usuario</i>	<i>Cercanía</i>	<i>Usuario</i>	<i>Cercanía</i>
<i>I. Arrimadas</i>	0,400582	<i>Íñigo Errejón</i>	0,323844
<i>P. Manglano</i>	0,381767	<i>Irene Montero</i>	0,322884
<i>Toni Cantó</i>	0,377760	<i>Marta Rovira</i>	0,319499
<i>J.L. Ábalos</i>	0,377248	<i>Oriol Junqueras</i>	0,317297
<i>X.G. Albiol</i>	0,374104	<i>Albert Rivera</i>	0,315245
<i>F. Páramo</i>	0,370376	<i>Arnaldo Otegui</i>	0,313339
<i>Pedro Sánchez</i>	0,367011	<i>Ximo Puig</i>	0,313337
<i>J.C. Girauta</i>	0,366861	<i>S. Abascal</i>	0,313313
<i>Ada Colau</i>	0,363687	<i>M. Carmena</i>	0,312251
<i>J.C. Monedero</i>	0,356746	<i>M.D. Cospedal</i>	0,307656
<i>Javier Maroto</i>	0,354024	<i>Quim Torra</i>	0,304908
<i>Gabriel Rufián</i>	0,354004	<i>F.M. Maíllo</i>	0,304848
<i>C. Puigdemont</i>	0,353046	<i>R. Hernando</i>	0,304204
<i>Pablo Casado</i>	0,340242	<i>Adriana Lastra</i>	0,301321
<i>R. Espinar</i>	0,337101	<i>Patxi López</i>	0,300187
<i>P. Echenique</i>	0,336251	<i>Rosa Díez</i>	0,289214
<i>A. Garzón</i>	0,334974	<i>Mariano Rajoy</i>	0,283134
<i>Fran Hervías</i>	0,330370	<i>E.G. Page</i>	0,275972
<i>Miquel Iceta</i>	0,330118	<i>Jordi Turull</i>	0,275226
<i>Pablo Iglesias</i>	0,329579	<i>Aitor Esteban</i>	0,270875
<i>J.M. Villegas</i>	0,327759	<i>Íñigo Urkullu</i>	0,262243
<i>Félix Álvarez</i>	0,327708	<i>Díaz/Santamaría</i>	0,25/0,23

En lo alto del ranking (tabla izquierda), observamos políticos que tenían ya un alto nivel de intermediación como Xavier Albiol o Inés Arrimadas y también otros con un grado de intermediación más bajo como Fernando de Páramo. Debemos destacar que la mayoría de los políticos que encontramos en lo alto de este ranking no se encuentran entre aquellos con un grado de salida mayor, por ello incidimos en la siguiente hipótesis: un político que responde o hace retuit a otros políticos que son centrales en la red y tienen un alto nivel de intermediación, puede tener un elevado nivel de cercanía sin ser “importante” en términos de intermediación o grado (Molina, 2001).

Si observamos los nodos de la Figura 3 podemos darnos cuenta de que aquellos con un tamaño mayor, y que por lo tanto tienen un mayor número de lazos con el grupo de estudio, se corresponden con aquellos políticos que lideraban el ranking de cercanía en la red global de sus interacciones. Inés Arrimadas es el usuario con mayor grado de salida en esta red con un total de

20, les siguen Fernando de Páramo con 15, Toni Cantó y Percival Manglano con un total de 14 interacciones de salida distintas. Si entendemos que los nodos más importantes para determinar la cercanía de la red global son los pertenecientes al grupo de estudio, puesto que son los que poseen interacciones de salida con el resto de nodos de la red, podremos comprender que cuantas más relaciones distintas tengamos en este grupo más cerca estaremos del resto de nodos.

En nuestro estudio observamos que Inés Arrimadas es un nodo importante de acuerdo a la mayoría de medidas de centralidad analizadas hasta el momento, gozando además de un grado de entrada elevado dentro de la red de interacciones entre políticos del grupo de estudio, sin embargo, la hipótesis sobre la calidad de las relaciones directas y su importancia en la determinación de la cercanía puede darse en el caso de Fernando de Páramo, Toni Cantó e incluso de Percival Manglano, ya que poseen un grado de entrada muy bajo. Estos últimos son políticos, sobre todo Páramo y Manglano, tienen una baja reciprocidad que puede provocar el fracaso en su intento de generar un debate con otros políticos y por lo tanto de conseguir seguidores o votantes para sus respectivos partidos. Continuando con los resultados locales vamos a mostrar una tabla con la centralidad del vector propio para el conjunto de los políticos analizados.

Tabla 5. Ranking basado en la centralidad del vector propio (total y grupo de estudio).

<i>Usuario</i>	<i>C_{VP}</i>	<i>Usuario</i>	<i>C_{VP}</i>
<i>Europapress</i>	1,0	<i>P. Echenique</i>	0,621
<i>Al Rojo Vivo</i>	0,992	<i>Ada Colau</i>	0,560
<i>eldiario.es</i>	0,927	<i>Pablo Iglesias</i>	0,548
<i>Tve 24h</i>	0,909	<i>Albert Rivera</i>	0,532
<i>El País</i>	0,854	<i>Inés Arrimadas</i>	0,498
<i>Espejo Público</i>	0,807	<i>X. Domenech</i>	0,467
<i>Sexta Noticias</i>	0,780	<i>Alberto Garzón</i>	0,465

En la tabla anterior tenemos dos rankings, el de la izquierda se corresponde con los siete usuarios con mayor centralidad basado en el vector propio y el de la derecha con aquellos políticos con una mayor centralidad basada en el mismo índice. Para el segundo ranking hemos tomado en cuenta todos los políticos que nos aparecen en la red, si bien, aquellos con una mayor centralidad se corresponden en su mayoría con los pertenecientes al grupo de estudio.

Los resultados sobre la centralidad del vector propio nos confirman de nuevo la importancia de los medios de comunicación y algunos espacios televisivos, éstos son utilizados por el conjunto de políticos para informarse y en muchas ocasiones para comunicar a sus seguidores noticias que perjudican a sus adversarios políticos o noticias que les favorecen a ellos mismos. Vuelven a destacar Europapress, “eldiario.es”, El País, el programa Al Rojo Vivo y el canal Tve 24h. Estos medios están conectados con un gran número de políticos centrales y por ello se ven afectados positivamente en este ranking. Por otro lado, entre los políticos con mayor índice de centralidad encontramos principalmente políticos pertenecientes a Unidos Podemos y Ciudadanos. A continuación, proseguimos con los resultados del nivel de excentricidad.

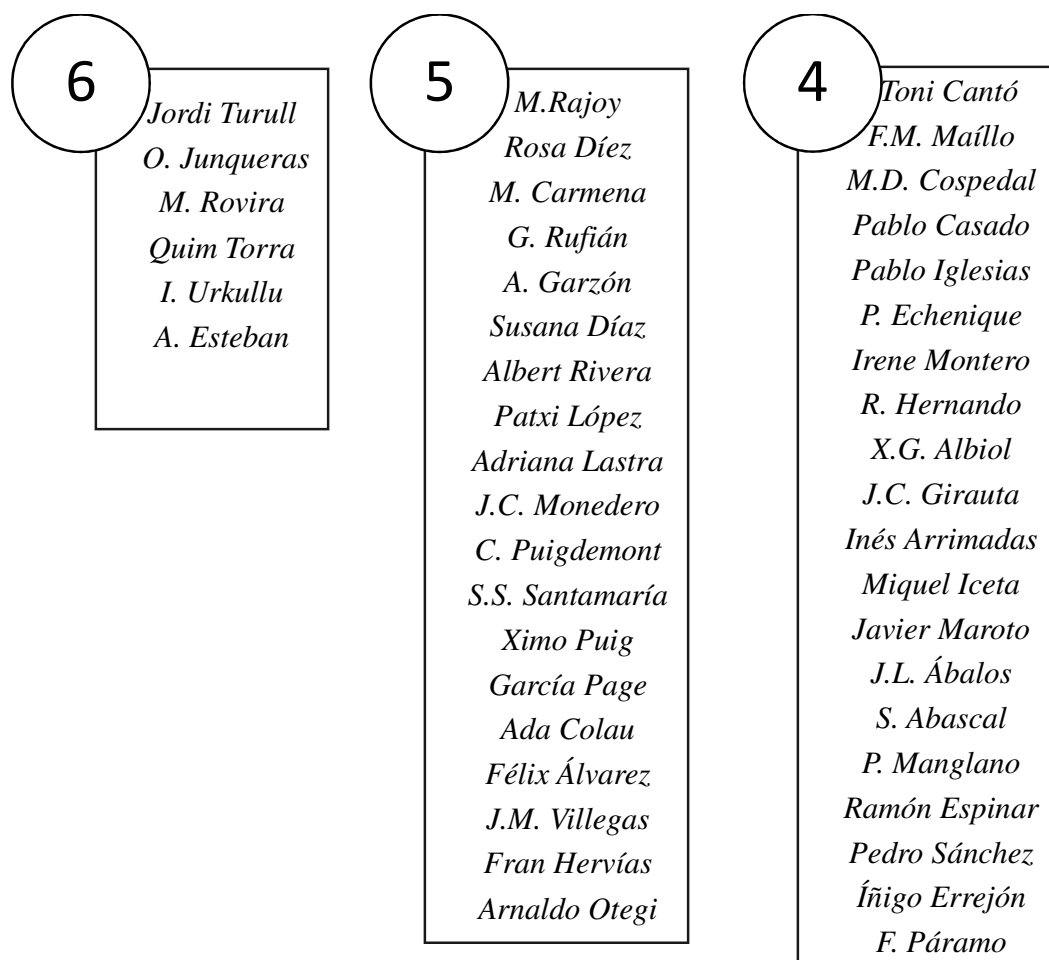


Figura 5. Nivel de excentricidad del grupo de estudio. Fuente: Elaboración propia con Gephi.

La principal conclusión que obtenemos sobre los resultados de la excentricidad es la posición periférica que ocupan el conjunto de políticos independentistas, pues el grupo formado por

aquellos que tienen una excentricidad de seis está formado completamente por políticos de este carácter, además, no encontramos ningún político perteneciente a un partido independentista en el grupo de políticos con menor excentricidad (más centrales de acuerdo a esta métrica). Por otro lado, destacamos que la mayoría de políticos poseen una excentricidad de 4, es decir, son 4 los enlaces que separan a estos usuarios respecto al usuario más alejado a ellos de la red. Además, también subrayamos la presencia en este grupo de los principales líderes de los partidos políticos más importantes, con la excepción del líder de Ciudadanos.

Para terminar con el análisis de la centralidad de los nodos vamos a realizar un filtrado en la herramienta Gephi. El objetivo es quedarnos con aquellos políticos del grupo de estudio que se sitúen entre los mejores en intermediación, cercanía y centralidad del vector propio a la vez. De esta forma podemos apreciar qué políticos de la red gozan de una centralidad en sentido conjunto y por lo tanto tienen un papel destacado en la red social.

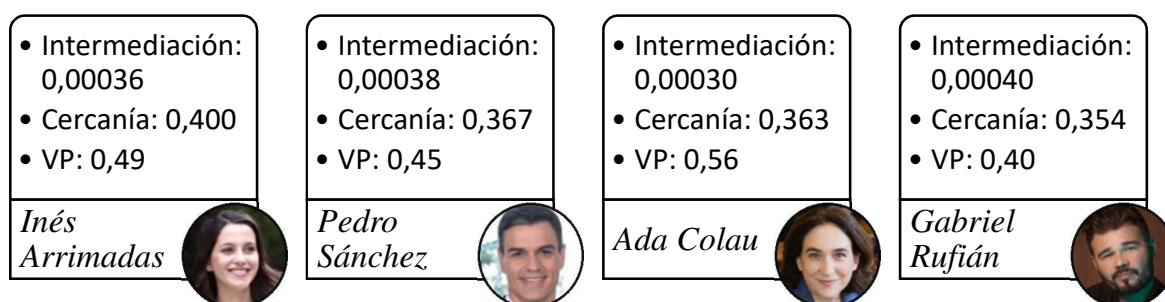


Figura 6. Políticos con mayor nivel de centralidad según intermediación, cercanía y centralidad del vector propio. Fuente: Elaboración propia con Gephi.

En la anterior figura, hemos calculado el valor del tercer cuartil en cada una de las métricas propuestas para el grupo de estudio y hemos filtrado aquellos políticos que se encuentran por encima de dicho valor en las tres medidas. El grado no lo hemos tenido en cuenta ya que las dos últimas medidas se encuentran correlacionadas con el grado, la cercanía con el de salida y la centralidad del vector propio con el de entrada.

Los resultados de la Figura 5 muestran los cuatro políticos (tres pertenecen a la política catalana) que gozan de una mayor centralidad y que por lo tanto tienen una elevada importancia en la red de interacciones, los cuatro políticos pertenecen a partidos distintos y pueden ejercer una gran

influencia debido a la repercusión de sus tuits dentro de la red. Son políticos que, independientemente de la positividad de sus interacciones de entrada, gozan de una gran difusión entre el resto de políticos y actúan como representantes políticos dentro de la misma.

Por último, vamos a mostrar una medida al margen de la centralidad. Ésta se corresponde con la media de retuits recibidos en cada tuit que realizan los miembros del grupo de estudio. Se trata de una medida sencilla que puede aportarnos datos reveladores sobre la difusión entre el resto de usuarios. Para su cálculo utilizamos R y adjuntamos el código en el Apéndice.

Tabla 6. Ranking según la media de los retuits recibidos.

<i>Nombre</i>	<i>Media RT</i>	<i>Nombre</i>	<i>Media RT</i>
<i>Carles Puigdemont</i>	4510	<i>Fernando Martínez Maíllo</i>	82
<i>Gabriel Rufián</i>	2374	<i>Aitor Esteban</i>	78
<i>Oriol Junqueras</i>	2125	<i>Ximo Puig</i>	76
<i>Alberto Garzón</i>	1454	<i>Rafael Hernando</i>	60
<i>Pablo Iglesias</i>	997	<i>Félix Álvarez</i>	54
<i>Inés Arrimadas</i>	977	<i>Patxi López</i>	45
<i>Quim Torra</i>	908	<i>Íñigo Urkullu</i>	41
<i>Santiago Abascal</i>	831	<i>Miquel Iceta</i>	38
<i>Jordi Turull i Negre</i>	826	<i>Emiliano García Page</i>	32
<i>Pablo Echenique</i>	822	<i>José Manuel Villegas</i>	31

En la tabla anterior tenemos los diez políticos con mayor y menor media de retuits. El ranking de los diez usuarios más retuiteados está liderado por políticos independentistas. También destaca la presencia de algunos miembros de Podemos. Por otro lado, tenemos los políticos cuya difusión es muy reducida. En este grupo situamos a políticos de diversos partidos y destacamos la ausencia de miembros de Podemos o políticos catalanes independentistas.

Una vez más, se ha vuelto a evidenciar la importancia del contexto catalán en la red social. Por otro lado, subrayamos la difusión que gozan los principales líderes de Unidos Podemos y destacamos en particular la presencia del líder de VOX (Santiago Abascal).

4.3 Detección de comunidades en la red

A lo largo del análisis global y local hemos observado varias figuras correspondientes a la red de interacciones del grupo de estudio. A menudo, hemos hecho referencia a la modularidad. A continuación, detallamos los resultados obtenidos en base a esta propiedad, donde el peso de las relaciones juega un papel fundamental.

Observando el color de los nodos de la Figura 2 podemos ver la red dividida en 13 comunidades, se puede apreciar que las comunidades que forman los políticos se corresponden en la mayoría de ocasiones con el partido político al que pertenecen. Si analizamos cada político del grupo de estudio podemos comprobar que salvo en algunos casos particulares, todos se encuentran clasificados de acuerdo a su partido político, esto es normal ya que la mayoría de sus interacciones (respuestas y retuits) son con miembros del propio partido. Cabe destacar algunas excepciones como el caso de Emiliano García-Page, el cual pertenece al PSOE pero queda aislado en una comunidad formada solamente por él y por aquellos con los que solo él tiene interacciones, o el caso de las alcaldesas de Madrid y Barcelona cuyos partidos regionales pertenecen a la formación Podemos pero aparecen en comunidades distintas para cada una de ellas. Estos sucesos desaparecen cuando tomamos en cuenta solamente las interacciones entre políticos (Figura 3). En dicha red el número de comunidades es menor debido a que el número de interacciones es sustancialmente menor. Además, encontramos que los partidos independentistas catalanes se agrupan en una única comunidad cuando observamos dicha red. Por otro lado, destaca también el caso de Percival Manglano. El concejal del Partido Popular por Madrid aparece en la misma comunidad que Santiago Abascal y VOX (parecen tener muchas interacciones con miembros comunes entre ellos), sin embargo, todos ellos aparecen en la misma comunidad que los miembros de Podemos en la Figura 4, este fenómeno tiene lugar debido al peso de las interacciones de estos políticos con miembros de Podemos en la red de interacciones locales. Más adelante explicaremos el motivo social de este suceso.

En el estudio de la centralidad de los nodos pudimos observar cómo el ranking del grado de entrada estaba formado en su mayoría por medios de comunicación y periodistas; éstos también aparecen entre los nodos con un alto grado de entrada con pesos y son una parte muy importante de la red.

Los medios son utilizados por el conjunto de los políticos en su actividad diaria. Gracias a la modularidad podemos situar estos nodos en la red y ver en qué comunidad se encuentran de acuerdo al uso que los políticos hacen de ellos.

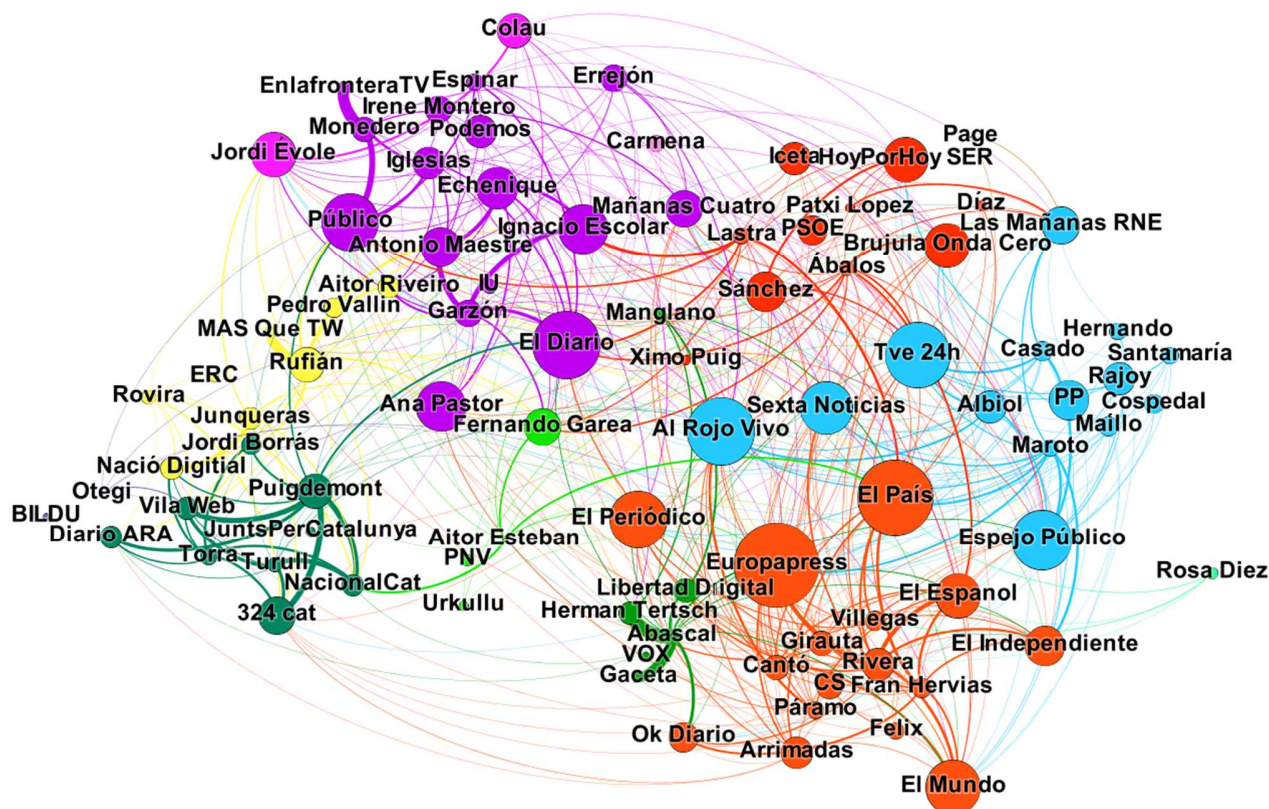


Figura 7. Red de interacciones con medios de comunicación. Fuente: Elaboración propia, Gephi.

En la red anterior tenemos situados aquellos medios y periodistas que contienen un mayor grado de entrada con pesos. El tamaño de los nodos se corresponde con el grado de entrada, las aristas que aparecen son las relaciones entre los miembros del grupo de estudio y los medios (para ello hemos realizado un filtrado con máscara en Gephi). Además el grosor de las aristas está en función del peso (intensidad de la relación). De este modo podemos establecer cuál es la preferencia de información de acuerdo a cada partido político y ver también a los periodistas que pueden ser más cercanos a unos u otros.

Los medios y contenidos de los mismos, así como algunos periodistas, aparecen distribuidos entre todas las comunidades. Observamos que en la comunidad formada por los políticos que pertenecen a Podemos (comunidad morada) aparecen medio digitales como Público o El Diario,

programas como Mañanas Cuatro o periodistas como Antonio Maestre, Ignacio Escolar y Ana Pastor. Destacamos el caso del periodista Jordi Évole que aparece en la comunidad de Ada Colau a pesar de tener relaciones con miembros del resto de comunidades. Esto sucede debido al gran peso de la relación con Ada, la cual ha retuiteado o respondido al periodista en 18 ocasiones; también se debe al número elevado de relaciones con el resto de miembros de Podemos. En la comunidad formada por los miembros del PSOE (comunidad en color rojo) encontramos solamente dos programas de radio, “Hoy por hoy” de la cadena SER y La Brújula de Onda Cero. La comunidad en la que encontramos más medios de comunicación es la formada por los miembros de Ciudadanos (comunidad naranja), en esta comunidad aparecen los medios tradicionales como El Mundo, El País, El Español, El Independiente, Ok Diario, El Periódico y Europapress. Además, dependiendo de dónde empiece la formación de comunidades el algoritmo de búsqueda (ésta es aleatoria), podemos obtener resultados un poco distintos y programas como Al Rojo Vivo y La Sexta Noticias pueden aparecer también en la comunidad formada por Ciudadanos, esto se debe a la cercanía que existe por parte de algunos medios con más de una comunidad. Sin embargo, tras diversas ejecuciones del algoritmo de comunidades, los resultados son idénticos salvo para algunos programas como los nombrados antes, además de Ok Diario que puede aparecer en la comunidad de VOX y Las Mañanas RNE que puede aparecer en la comunidad del PSOE. En la comunidad del PP (comunidad en azul) encontramos el programa Espejo Público o el canal 24h de TVE. Por último, encontramos medios como Libertad Digital o La Gaceta en la comunidad de VOX (nodos en color verde) y medios de Cataluña en las comunidades de los partidos independentistas (color amarillo para ERC y verde menta para el PDeCAT) como “324 Cat” o Nació Digital.

Hemos clasificado, de acuerdo a la intensidad de las relaciones por parte de los políticos del grupo de estudio, a los medios de comunicación y periodistas más importantes de la red, llegando a la conclusión de que los miembros de los distintos partidos utilizan las fuentes de información más cercanas a su ideología política, económica y social ya que en la mayoría de los casos las relaciones son retuits. En la red global aparecen muchos más medios y programas de los que hemos mostrado, pero éstos tienen un grado de entrada con pesos inferior a 60, el cual es el tope que hemos utilizado en la selección. Otros medios como el ABC o La Razón aparecen respectivamente en las comunidades de Ciudadanos y PP.

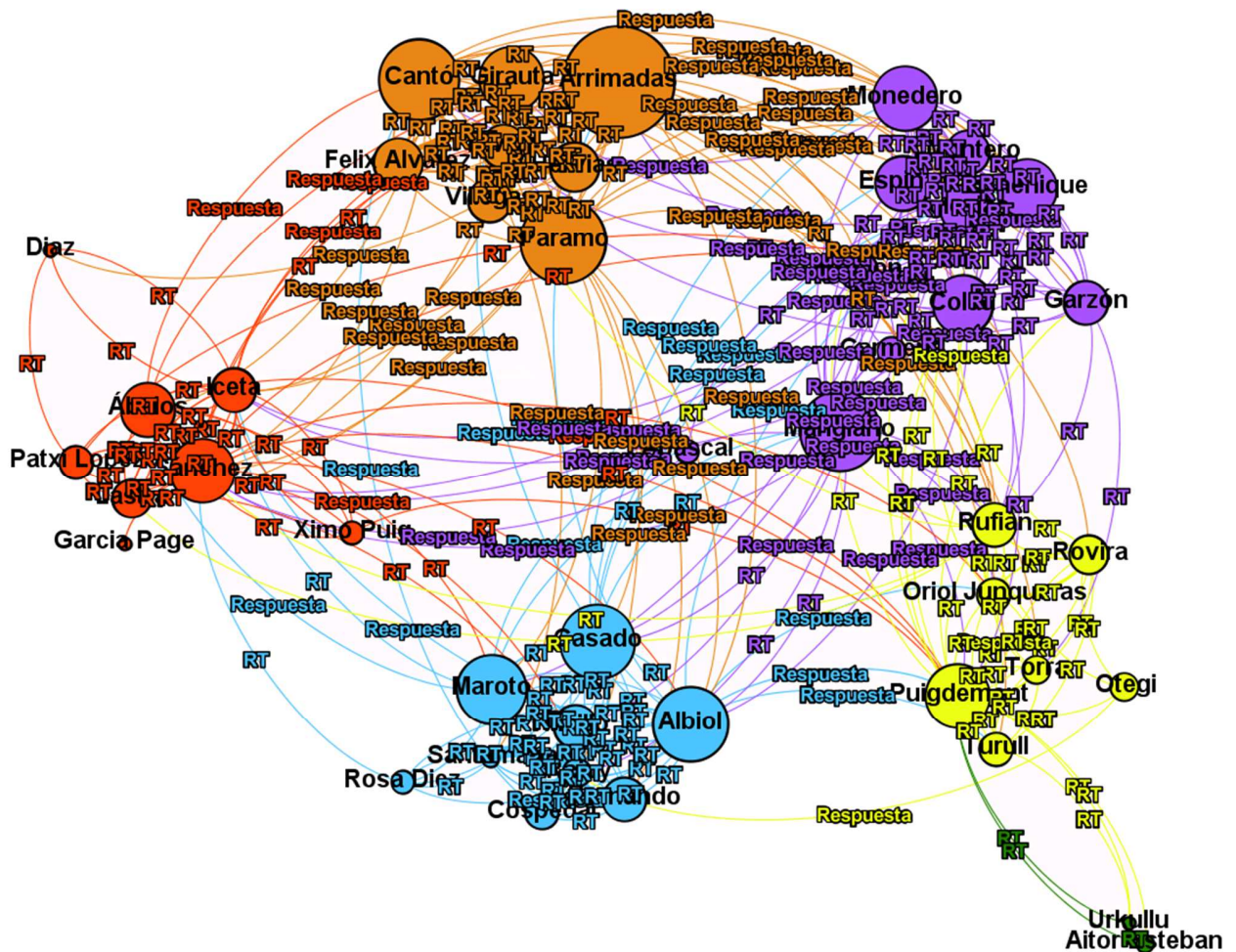


Figura 8. Red de interacciones entre miembros del grupo de estudio (tipo de relación). Fuente: Elaboración propia con Gephi

En la figura anterior tenemos la tipología de las conexiones, el objetivo es comprender los motivos sociales de algunas relaciones particulares que nos hemos encontrado a lo largo del estudio.

Hay varias conclusiones que podemos obtener en base a la tipología relacional. La primera se corresponde con la concentración de los retuits que se produce dentro de cada comunidad. Los políticos pertenecientes a una misma comunidad (en nuestro caso la mayoría forman el partido político al que pertenecen) tienden a apoyarse entre sí y no suelen poner en duda las afirmaciones de sus compañeros en la red social, de ahí que las respuestas se concentren en las relaciones intercomunitarias. Observamos que tanto el Partido Popular como Ciudadanos, los cuales se engloban en las comunidades azul y naranja, son los que más respuestas tienen con el resto de

partidos, destacando la comunidad naranja como la que más responde. Tenemos que aclarar, que en líneas generales, las respuestas entre miembros de partidos políticos con ideologías opuestas suelen ser de carácter crítico, además, éstas generan un debate que puede atraer a seguidores dependiendo de la calidad del mismo. Por otro lado, podemos ver algunos retuits de algunos miembros de Podemos hacia la comunidad independentista catalana (lo que implica una cercanía con esta comunidad que puede deberse a los últimos sucesos en tierras catalanas), así mismo, no poseen ningún otro retuit con otra comunidad ya que los que podemos apreciar en la red en favor del PP son los realizados por Percival Manglano, el cual ha sido agrupado en la comunidad morada de acuerdo a su gran número de interacciones con miembros de este grupo (todas son respuestas). Este suceso no nos preocupa, ya que como comentábamos anteriormente solo tiene lugar en esta red. Una vez observamos la red global podemos comprender que Percival Manglano se relaciona con miembros de una ideología más afín a la comunidad que engloba a Abascal y VOX.

Otra de las conclusiones que encontramos en la Figura 8 está relacionada con el comportamiento que exhiben las distintas comunidades, partidos como el PSOE o Podemos se comportan de una forma menos confrontativa con el resto de partidos, sobre todo el Partido Socialista. Por otro lado, partidos como Ciudadanos o el PP buscan constatemente la confrontación directa. Es necesario aclarar que en muchas ocasiones este comportamiento se debe al comportamiento de algunos miembros de la comunidad en concreto y en otros casos se tratan de comportamientos esporádicos, ya que en líneas generales las interacciones entre miembros de distintos partidos es reducida (de acuerdo al número máximo de tuits por usuario que disponemos para el análisis).

Por último, vemos que las relaciones dentro de la comunidad formada por políticos independentistas catalanes procedentes de distintos partidos son mayormente relaciones de apoyo, apenas existe una única respuesta de Torra a Junqueras. A pesar de pertenecer a partidos con ideología económica y social distinta, la independencia forma una unión en la red social.

Para terminar con los resultados en este apartado de detección de comunidades, vamos a justificar la aparición de Percival Manglano en la comunidad formada por los miembros de Unidos Podemos.

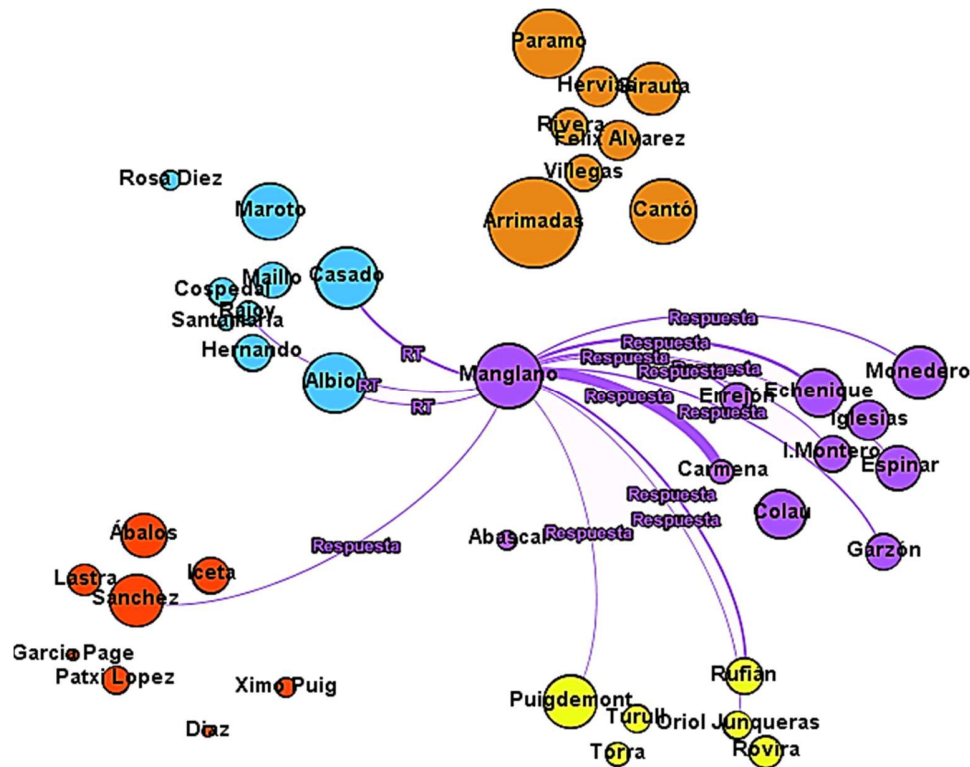


Figura 9. Red de interacciones de Percival Manglano con el resto de miembros del grupo de estudio (tipo de relación). Fuente: Elaboración propia con Gephi

En la figura anterior, comprobamos que Manglano es equivocadamente introducido en la comunidad de Podemos debido a la cantidad de respuestas que realiza a sus miembros, en particular a Manuela Carmena como consecuencia del papel que juega como alcaldesa de Madrid. También encontramos otros casos como el de Rosa Díez que aparece en la comunidad del PP, si bien, el número de interacciones que posee es tan reducido que el resultado obtenido al igual que en el caso de Santiago Abascal no es significativo.

4.4 Análisis del sentimiento político en varios medios de comunicación

En este apartado vamos a analizar el sentimiento político de los titulares correspondientes a noticias de ámbito político, más en concreto de noticias correspondientes a alguno de los cuatro grandes partidos políticos a nivel nacional en el estado español.

Cabe mencionar que la recogida de los titulares se corresponde con un periodo aproximado de tres meses (entre Julio y Octubre de 2018), si bien, no estrictamente continuado. Además el número de titulares recogidos entre cada medio de comunicación es distinto, ya que depende de lo que ofrecen.

Una vez obtenidos los titulares, lo primero que hemos realizado es su clasificación mediante un diccionario de términos relacionados con alguno de los cuatro partidos. Además, hemos intentado separar aquellos titulares que se corresponden con sucesos de otros que se corresponden con lo que dice o manifiesta un miembro de un partido. Cabe destacar que tras su clasificación se han podido repasar manualmente para corregir algunos errores graves que pudieran aparecer, si bien, alrededor del 95% se encontraban bien clasificados de acuerdo al partido que pertenece.

Tras clasificar los titulares según el partido político al que hacen referencia, hemos realizado la clasificación de acuerdo al sentimiento. Para ello, hemos utilizado la API de Meaning Cloud. Cabe destacar que la herramienta de análisis de sentimiento de Meaning Cloud tiene una precisión aceptable, ya que realiza el análisis por partes, tomando en cuenta también el sentido y la posible subjetividad de la oración.

Antes de comenzar a comentar los resultados de la Tabla 6 debemos realizar algunas consideraciones. Los porcentajes observados hacen referencia a la proporción sobre el total de titulares que se refieren a cada partido. Si sumamos en horizontal obtenemos el 100% (en las tablas se han omitido los decimales). Para el conjunto de medios, el total de titulares que hacen referencia al PSOE es superior al resto de partidos. El PP también destaca situándose cerca del PSOE. Por último, como era de esperar, aquellos partidos para los que se ofrecen menos noticias en los medios se corresponden con Ciudadanos y Unidos Podemos. Además, el total de titulares analizados para cada medio son: 593 para ABC, 397 para El Mundo, 354 para El País, 846 para “eldiario.es” y 733 para Público.

A continuación, vamos a exponer y comentar los resultados obtenidos en el análisis de sentimientos.

Tabla 7. Análisis de sentimiento de titulares de noticias de “ABC”.

<i>Partido</i>	<i>Análisis de sentimiento (ABC)</i>					
	Positividad		Neutralidad		Negatividad	
	Sucesos	Manifestaciones	Sucesos	Manifestaciones	Sucesos	Manifestaciones
<i>PSOE</i>	11%	7%	28%	7%	38%	6%
<i>PP</i>	18%	7%	35%	9%	10%	19%
<i>C's</i>	13%	2%	30%	5%	22%	25%
<i>UP</i>	9%	6%	26%	6%	37%	13%

La primera conclusión que obtenemos de la Tabla 6 se corresponde con los altos porcentajes de neutralidad obtenidos para titulares de sucesos que afectan a los distintos partidos. Destaca principalmente el PP, donde el 35% de los titulares de sucesos que afectan al partido o a sus miembros son de carácter objetivo, es decir, no poseen un sentimiento polarizado. Por otro lado, si observamos los porcentajes de titulares que se corresponden con sucesos que denotan un sentimiento positivo, podemos apreciar que éstos son inferiores para los partidos políticos de ideología progresista. Del mismo modo, el porcentaje de titulares sobre sucesos polarizados negativamente es superior en partidos de la misma índole. De acuerdo a las manifestaciones que hacen los políticos, subrayamos los altos niveles de negatividad en los partidos de ideología conservadora. Durante los tres meses del análisis, estos partidos se sitúan en la oposición y el tono de sus mensajes suele ser más elevado.

Tabla 8. Análisis de sentimiento de titulares sobre noticias de “El Mundo”.

<i>Partido</i>	<i>Análisis de sentimiento (El Mundo)</i>					
	Positividad		Neutralidad		Negatividad	
	Sucesos	Manifestaciones	Sucesos	Manifestaciones	Sucesos	Manifestaciones
<i>PSOE</i>	10%	11%	24%	4%	35%	13%
<i>PP</i>	18%	3%	27%	8%	21%	20%
<i>C's</i>	7%	7%	46%	7%	7%	23%
<i>UP</i>	2%	2%	8%	4%	55%	27%

En los resultados obtenidos para “El Mundo”, subrayamos en general los altos niveles de negatividad en titulares de sucesos, destacando Podemos y PSOE como los partidos con niveles más elevados. El caso de Ciudadanos nos llama la atención. Este partido se sitúa con el mayor nivel de neutralidad para los titulares referentes a sucesos y también con el menor nivel de negatividad. De acuerdo a los titulares referentes a manifestaciones o mensajes de políticos,

observamos un nivel de negatividad que sigue la línea observada para ABC, si bien, destaca el caso de Podemos con un elevado porcentaje por encima del resto.

Tabla 9. Análisis de sentimiento de titulares sobre noticias de “El País”.

<i>Partido</i>	<i>Análisis de sentimiento (El País)</i>					
	Positividad		Neutralidad		Negatividad	
	Sucesos	Manifestaciones	Sucesos	Manifestaciones	Sucesos	Manifestaciones
<i>PSOE</i>	20%	13%	39%	6%	10%	9%
<i>PP</i>	8%	5%	40%	5%	27%	11%
<i>C,s</i>	11%	5%	38%	8%	11%	25%
<i>UP</i>	9%	9%	28%	0%	31%	21%

Observando el análisis de El País, vemos que el nivel de neutralidad general en los sucesos es muy elevado. En el porcentaje de titulares positivos y negativos destaca el PSOE, que tanto en sucesos como manifestaciones, obtiene los mejores resultados. Por otro lado, el PP y Podemos reciben los peores resultados en sucesos.

Tabla 10. Análisis de sentimiento de titulares sobre noticias de “eldiario.es”.

<i>Partido</i>	<i>Análisis de sentimiento (eldiario.es)</i>					
	Positividad		Neutralidad		Negatividad	
	Sucesos	Manifestaciones	Sucesos	Manifestaciones	Sucesos	Manifestaciones
<i>PSOE</i>	7%	8%	41%	29%	5%	7%
<i>PP</i>	4%	4%	31%	25%	16%	16%
<i>C's</i>	1%	7%	12%	37%	11%	27%
<i>UP</i>	4%	8%	27%	38%	5%	15%

Atendiendo a los resultados de “eldiario.es”, podemos afirmar que los partidos de ideología conservadora son los que peores resultados arrojan, tanto en titulares sobre sucesos como en manifestaciones, si bien, debemos tomar en cuenta que el conjunto de porcentajes negativos y positivos se sitúan en un rango inferior respecto al resto de medios. Subrayamos los elevados niveles de neutralidad para los mensajes que transmiten el conjunto de partidos. Por último, destacamos el caso del PSOE, que vuelve a obtener los mejores resultados.

Tabla 11. Análisis de sentimiento de titulares sobre noticias de “Público”.

<i>Partido</i>	<i>Análisis de sentimiento (Público)</i>					
	Positividad		Neutralidad		Negatividad	
	Sucesos	Manifestaciones	Sucesos	Manifestaciones	Sucesos	Manifestaciones
<i>PSOE</i>	6%	3%	53%	25%	7%	3%
<i>PP</i>	3%	1%	41%	21%	23%	8%
<i>C's</i>	3%	0%	27%	14%	34%	20%
<i>UP</i>	6%	4%	38%	35%	7%	8%

Por último, tenemos los resultados para Público. Este diario presenta unos niveles de neutralidad muy alta. Los porcentajes de titulares objetivos sin ninguna polaridad son elevados tanto para sucesos como manifestaciones. Los porcentajes de noticias positivas se sitúan en un rango muy reducido. Si observamos los porcentajes de titulares negativos sobre sucesos, apreciamos que los partidos conservadores salen más perjudicados, en especial Ciudadanos, que posee los peores resultados tanto en porcentaje de titulares positivos como negativos y neutrales.

En líneas generales, podemos comentar varios aspectos respecto al conjunto de los medios analizados. Los medios tienen una elevada neutralidad, ya que los porcentajes de titulares neutrales u objetivos referentes a sucesos son los más altos. Público y “eldiario.es” son los medios más objetivos. Cabe destacar que la polarización realizada en un principio a los medios, con el objetivo de establecer si tenían lugar las hipótesis de cercanía, se cumple en líneas generales. ABC y El Mundo arrojan mejores resultados para los partidos conservadores, Público y “eldiario.es” proveen mejores resultados para el lado progresista, si bien, éstos son más neutrales que los dos anteriores. En el caso de El País, los resultados son dispares entre ideologías, arrojando los mejores resultados para Ciudadanos y especialmente el PSOE. Por último, atendiendo a los titulares que se corresponden con manifestaciones que hacen los políticos o los partidos, y que por lo tanto representan el sentir de cada partido respecto a la actualidad política y social, podemos establecer que durante estos primeros meses de gobierno los mensajes que llegan del PSOE son los más positivos, por el contrario, el conjunto de partidos de la oposición y UP parecen tener un tono más negativo en sus mensajes.

5 Conclusiones y posibles extensiones del trabajo

A lo largo del trabajo hemos obtenido varias conclusiones sobre el comportamiento de los políticos españoles en la red social Twitter. La primera se corresponde con la importancia que tienen los medios de comunicación para el conjunto de los políticos en dicha red. Todos los políticos analizados utilizan dichos medios como fuentes de información y difunden su contenido. Entre estos medios destacan Europapress y El País, siendo utilizados por la mayoría de políticos. Por otro lado, destacamos los políticos que más interactúan con el resto de usuarios, donde predominan políticos independentistas, despuntando por encima de todos Gabriel Rufián. También subrayamos, en líneas generales, la importancia de la política catalana en el conjunto de la red de acuerdo a su presencia en las distintas métricas. En las interacciones con el resto de políticos en la red destacan los pertenecientes a Ciudadanos, donde sobresale Inés Arrimadas. Tanto la interacción con usuarios como con el resto de políticos puede atraer el interés de otros usuarios. Si atendemos al conjunto de métricas locales, los políticos más importantes en la red son: Inés Arrimadas, Pedro Sánchez, Ada Colau y Gabriel Rufián.

En la detección de comunidades hemos destacado la situación de los medios de comunicación en la red, así como la tipología de las interacciones y el comportamiento de los partidos.

Entre los medios de comunicación más utilizados aparecen algunos como “eldiario.es” o Público en la comunidad de Unidos Podemos, otros como Europapress, El Mundo o El País en la comunidad de Ciudadanos y algunos programas televisivos como Espejo Público en la comunidad formada por miembros del Partido Popular.

De acuerdo al comportamiento de los partidos políticos con el resto de partidos, destacamos una concentración de la actividad. La mayoría de partidos políticos se relacionan muy poco con el resto de partidos o en algunos casos, si hay una relación, se debe a casos puntuales de algún miembro del partido. El único partido que destaca positivamente en este aspecto es Ciudadanos.

Por último, respecto al último análisis realizado en el trabajo destacamos tres aspectos: la elevada neutralidad general en los medios analizados, la cercanía ideológica de los distintos medios con los partidos políticos que se situaban cerca en la red y la polarización entre el PSOE y el resto de partidos de acuerdo al sentimiento que transmiten durante los primeros meses de gobierno socialista.

Como extensiones del trabajo vamos a plantear varios puntos:

- Realizar el análisis de las interacciones de los políticos aumentando el número de políticos en el grupo de estudio para obtener una red más representativa.
- Extender el análisis a otras redes sociales como Facebook.
- Analizar el sentimiento de noticias completas en lugar de titulares, si bien, el número de éstas podría ser más reducido.
- Realizar el análisis de sentimientos utilizando algún algoritmo de aprendizaje supervisado (recomendamos Maquinas de Soporte Vectorial o Naïve Bayes), junto con una muestra de noticias ya clasificadas como conjunto de entrenamiento (nosotros intentamos realizarlo con un diccionario de sentimientos y el algoritmo Naïve Bayes pero los resultados no llegaban a ser aceptables debido a la pobreza del diccionario).

Referencias

- BARABÁSI, A.-L., 2016. Graph theory. En: BARABÁSI, A.-L., *Network Science*. Glasgow: Cambridge University Press. pp.42-71. ISBN 978-1-107-07626-6
- BLONDEL, V.D., et al; 2008. Fast unfolding communities in large networks. En: *Journal of Statistical Mechanics Theory Experiment*. Marc Mézard. Vol. 2008.
- BONACICH, P., 1987. Power and Centrality: A Family of Measures. En: *American Journal of Sociology*. Vol. 92. Chicago. No. 5, pp. 1170-1182
- CAMPOS-DOMINGUEZ, E., 2017. Twitter y la comunicación política. En: *El profesional de la información*. Vol 26. no 5, pp. 785-793. eISSN: 1699-2407
- CHAVES-MONTERO, A., et al; 2017. Uso, efectividad y alcance de la comunicación política en las redes sociales. En: CHAVES-MONTERO, A., ET AL; *Comunicación Política y Redes Sociales*. Sevilla: Alfonso Chaves-Montero. pp. 1-32. ISBN 978-84-17270-09-4
- FREEMAN, L.C., 1978. Centrality in social networks conceptual clarification. En: *Social networks*. Netherland: Elsevier Sequoia. Vol 1, pp. 215-239
- GODSIL, C. Y ROYLE, G., 2001. *Algebraic Graph Theory*. New York: Springer-Verlag. ISBN 978-0-387-95220-8
- HANNEMAN, R.A., RIDDLE, M., 2005. *Introduction to social networks methods* [en línea]. California: Riverside [consulta: 20 de septiembre 2018]. Disponible en: http://faculty.ucr.edu/~hanneman/nettext/Introduction_to_Social_Network_Methods.pdf
- IZQUIERDO, L.R. Y HANNEMAN, R.A., 2006. *Introduction to the formal analysis of social networks using mathematica* [en línea]. Burgos [consulta: 30 de septiembre 2018]. Disponible en: http://library.wolfram.com/infocenter/MathSource/6638/Izquierdo_Hanneman.pdf
- MOLINA, J.L., 2001. *El análisis de redes sociales Una introducción*. Barcelona: Edicions Bellaterra. ISBN 84-7290-160-2

- ROSEN, K.H., et al; 2000. *Handbook of discrete and combinatorial mathematics*. Florida. ISBN 0-8493-0149-1
- RUOHONEN, K., 2013. "Graph theory". En: *Keijo Ruohonen* [en línea], [consulta 8 de octubre 2018]. Disponible en: http://math.tut.fi/~ruohonen/GT_English.pdf
- WASSERMAN, S. Y FAUST, K., 1994. *Social network analysis: methods and applications*. New York: Cambridge University Press. ISBN 0 521 38707 8
- WHITE, A., 2017. El periodismo ético vuelve a primera plana. En: *El correo de la UNESCO*. Paris. no 2, pp. 7-11. ISSN 2220-2307

Apéndice

Código correspondiente a la extracción y depuración de los tuits con R:

```
#Limpiamos de variables nuestro entorno.
rm(list = ls())

#Cargamos la librería que nos va a permitir usar las funciones de extracción de tuits.
library(twitter)

#Cargamos los datos CSV en una variable llamada Politicos.
Politicos=read.csv("C:/Users/Luis/Desktop/ListaPoliticos.csv",sep=";",encoding='UTF-8')

#Nos registramos en Twitter para desarrolladores y guardamos las claves.
consumer_key="9jbfwmwdHg017vt9wuNjnyr1"
consumer_secret="mM3FcbZgyCIhQV8i5ENk64LGQBEyUqAmchcNsc5Nq1wyebaAwa"
access_token="414064333-2Ns1rgKjP3CyifZ9HbKGZ32rQ86SvPDXkq4H1dNV"
access_secret="wjvynxd2cSqnCU2niF3fDofwCN9LSzTTE8rYakuv8zjs0"

#Iniciamos usando las claves obtenidas.
setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secret)

#Creamos un panel de datos vacio.
dftp=data.frame(stringsAsFactors=FALSE)

#Vamos completando el panel de datos con los tuits de cada político.
#Para cada timeline en formato lista lo pasamos a panel de datos y lo unimos al panel vacío.
for(i in Politicos$Twitter) {
  timeline=userTimeline(i, n=3000, excludeReplies=FALSE,includeRts=T)
  dftp=rbind(dftp,twListToDF(timeline))}

#Filtramos para mantener solo las respuestas de los políticos recogidos.
primerfiltronormal=subset.data.frame(dftp,dftp$replyToSN!=dftp$screenName)

#Filtramos de nuevo para mantener las que son entre políticos.
primerfiltro=subset.data.frame(dftp,dftp$replyToSN %in% Politicos$Twitter)

#Filtramos los RT y separamos el texto para saber a quien están retuiteando.
#Además, creamos una nueva variable para guardar esos nombres en el panel de datos "whoisRT".
segundofiltro1parte=subset.data.frame(dftp,dftp$isRetweet==TRUE)

whoisRT=list()
for (i in segundofiltro1parte$text) {
  separar=unlist(strsplit(i,split="@"))
  separar2=unlist(strsplit(separar[2],split=":"))
  whoisRT=append(whoisRT,separar2[1])
}

for (i in 1:length(whoisRT)){
  segundofiltro1parte$whoisRT[i]=whoisRT[i]
}

#Mantenemos solo aquellos que no se hacen RT a sí mismos.
tercerfiltro=subset.data.frame(segundofiltro1parte,
                               segundofiltro1parte$screenName!=segundofiltro1parte$whoisRT)
```

```

#Unificamos los datos en un único data frame.
#Para ello, ya que el número de columnas no coincide, usamos el paquete gtools.
install.packages("gtools")
library(gtools)

datos=smartbind(primerfiltronormal,tercerfiltro)
datos0=smartbind(primerfiltro,tercerfiltro)

#Ahora vamos a filtrar solo los tuits que son RT o respuesta entre los políticos seleccionados.
datos1=subset.data.frame(datos0, (datos0$whoisRT %in% Políticos$Twitter) |
                             (datos0$replyToSN!=datos0$screenName))

#Creamos la tabla de enlaces para gephi y la guardamos en formato csv.
#Dependiendo de si queremos las interacciones globales o solo entre políticos.
#Usaremos "datos" o "datos1" según necesitemos.
onlyresponse=subset.data.frame(datos,datos$isRetweet==FALSE)
linksresponse=data.frame(Source=onlyresponse$screenName,Target=onlyresponse$replyToSN,
                          weight=1,Label="Respuesta",Type="Directed")

onlyRT=subset.data.frame(datos,datos$isRetweet==TRUE)
linkRT=data.frame(Source=onlyRT$screenName,Target=onlyRT$whoisRT,
                   weight=1,Label="RT",Type="Directed")

link=rbind(linksresponse,linkRT)
write.csv(link,"Global Links.csv",row.names=FALSE)

```

Código para extraer la media de los RT recibidos en R:

```

#Vamos a establecer ahora un ranking de influencia basado en una media de los RT que reciben.
#Para ello necesitamos un panel de datos que contenga solo los tweet que no son RT ni respuestas.
onlyTweet=subset.data.frame(dftp,(dftp$isRetweet==FALSE)&(is.na(dftp$replyToSN)))

Means=list()

for (x in Políticos$Twitter){
  framex=subset.data.frame(onlyTweet,onlyTweet$screenName==x)
  Means=append(Means,mean(framex$retweetCount))
}

Rank=data.frame(Politico=Políticos$X.U.FEFF.Nombre)

for (i in 1:length(Means)){
  Rank$MediaRT[i]=round(as.numeric(Means[i]),digits=0)
}

```

Extracción de titulares de eldiario.es con Python:

```
import requests
from bs4 import BeautifulSoup
import pandas as pd
import time

ElDiario=pd.DataFrame()
#Creamos un panel de datos vacío para guardar los titulares.

#Diario.es

page=requests.get("https://www.eldiario.es/politica/")
#Solicitud a la página del medio.

soup=BeautifulSoup(page.content,'html.parser')
#Extraemos los datos html.

politica=soup.find_all("li")
#Buscamos para encontrar los titulares.

fecha=[]
Noticia=[]

for x in politica:
    #Guardamos los titulares y la fecha en listas.
    try:
        Noticia.append(x.find("h2").find("a")["title"])
        fecha.append(time.strftime("%d/%m/%y"))
    except:
        continue

ElDiario["Time"]=fecha
ElDiario["Noticia"]=Noticia
#Creamos dos columnas en el panel de datos para guardar los titulares y la fecha.

ayer=pd.read_csv("Eldiario.csv",sep="," ,encoding='UTF-8',index_col="Indice")
hoy=ayer.append(ElDiario)
#Vamos fusionando los datos que ya teníamos con los recogidos hoy.

hoy.to_csv("Eldiario.csv",encoding='UTF-8',index_label="Indice")
#Guardamos en csv la fusión de lo acumulado y las nuevas noticias de hoy.

#El código se ejecuta cada día para recoger los titulares.
```

Extracción de titulares de El País con Python:

```

import requests
from bs4 import BeautifulSoup
import pandas as pd
import time

ElPais=pd.DataFrame()
#Creamos un panel de datos vacío para guardar los titulares.

#Pais.es

page=requests.get("https://politica.elpais.com/")
#Solicitud a la página del medio.

soup=BeautifulSoup(page.content,'html.parser')
#Extraemos los datos html.

secciones=soup.find_all("section")
#Buscamos las secciones que contienen los titulares.

fecha=[]
Noticia=[]

for v in secciones:
#Guardamos los titulares y la fecha en listas.
    for r in v.find_all("article"):
        try:
            Noticia.append(r.find("h2").find("a").get_text())
            fecha.append(time.strftime("%d/%m/%y"))
        except:
            continue

ElPais["Time"]=fecha
ElPais["Noticia"]=Noticia
#Creamos dos columnas en el panel de datos para guardar los titulares y la fecha.

ayer=pd.read_csv("ElPais.csv",sep=";",encoding='UTF-8',engine="python")
hoy=ayer.append(ElPais)
#Vamos fusionando los datos que ya teníamos con los recogidos hoy.

ElPais.to_csv("ElPais.csv",encoding='UTF-8',index_label="Indice")
#Guardamos en csv la fusión de lo acumulado y las nuevas noticias de hoy.

#El código se ejecuta cada día para recoger los titulares.

```

Extracción de titulares de El Mundo con Python:

```

import requests
from bs4 import BeautifulSoup
import pandas as pd
import time

ElMundo=pd.DataFrame()
#Creamos un panel de datos vacío para guardar los titulares.

fecha=[]
Noticia=[]

#El Mundo Espana

pagel=requests.get("http://www.elmundo.es/espana.html")
#Solicitud a la página del medio.

soulp1=BeautifulSoup(pagel.content, 'html.parser')
#Extraemos los datos html.

secciones1=soulp1.find_all("section")
#Buscamos las secciones que contienen los titulares.

contador1=0
for x in secciones1:
    #Guardamos los titulares y la fecha en listas.
    if contador1==3:
        break
    elif x["class"][0]==u'row-layout' and len(x["class"])==1:
        contador1+=1
        try:
            articulos=x.find_all("article")
            for i in articulos:
                try:
                    Noticia.append(i.find("h2").find("a").get_text())
                    fecha.append(time.strftime("%d/%m/%y"))
                except:
                    continue
        except:
            print "Hubo un apartado sin poder hacer Scrap"

ElMundo["Time"]=fecha
ElMundo["Noticia"]=Noticia
#Creamos dos columnas en el panel de datos para guardar los titulares y la fecha.

ayer=pd.read_csv("ElMundo.csv",sep=";",encoding='UTF-8',index_col="Indice")
hoy=ayer.append(ElMundo)
#Vamos fusionando los datos que ya teniamos con los recogidos hoy.

hoy.to_csv("ElMundo.csv",encoding='UTF-8',index_label="Indice")
#Guardamos en csv la fusion de lo acumulado y las nuevas noticias de hoy.

#El código se ejecuta cada día para recoger los titulares.

```

Extracción de titulares de ABC con Python:

```

import requests
from bs4 import BeautifulSoup
import pandas as pd
import time

ABC=pd.DataFrame()
#Creamos un panel de datos vacío para guardar los titulares.

#ABC Espana

pagel=requests.get("https://www.abc.es/espana/")
#Solicitud a la página del medio.

soulpl=BeautifulSoup(pagel.content,'html.parser')
#Extraemos los datos html.

destacadaespana=soulpl.find("main").find("span", class_=u"caja-portada")
espana=soulpl.find("main").find("section")
#Buscamos las secciones de los titulares.

try:
    articulos=destacadaespana.find_all("article")
    for i in articulos:
        #Guardamos los titulares y la fecha en listas.
        try:
            Noticia.append(i.find("h3").find("a").get_text())
            fecha.append(time.strftime("%d/%m/%y"))
        except:
            continue
except:
    print "Fallo al buscar articulos"

try:
    articulos=espana.find_all("article")
    for i in articulos:
        #Guardamos los titulares y la fecha en listas.
        try:
            Noticia.append(i.find("h3").find("a").get_text())
            fecha.append(time.strftime("%d/%m/%y"))
        except:
            continue
except:
    print "Fallo al buscar articulos"

ABC["Time"]=fecha
ABC["Noticia"]=Noticia
#Creamos dos columnas en el panel de datos para guardar los titulares y la fecha.

ayer=pd.read_csv("ABC.csv",sep="," ,encoding='UTF-8',index_col="Indice")
hoy=ayer.append(ABC)
#Vamos fusionando los datos que ya teniamos con los recogidos hoy.

hoy.to_csv("ABC.csv",encoding='UTF-8',index_label="Indice")

```

Extracción de titulares de Público con Python:

```

import requests
from bs4 import BeautifulSoup
import pandas as pd
import time

Publico=pd.DataFrame()
#Creamos un panel de datos vacío para guardar los titulares.

#Publico

page=requests.get("http://www.publico.es/politica")
#Solicitud a la página del medio.

soup=BeautifulSoup(page.content,'html.parser')
#Extraemos los datos html.

articulos=soup.find_all("div", class_="listing-item")
#Buscamos la sección de los titulares.

fecha=[]
Noticia=[]

for x in articulos:
    #Guardamos los titulares y la fecha en listas.
    for r in x.find_all("div", class_="listing-title"):
        try:
            Noticia.append(r.find("a").get_text())
            fecha.append(time.strftime("%d/%m/%y"))
        except:
            continue

Publico["Time"]=fecha
Publico["Noticia"]=Noticia
#Creamos dos columnas en el panel de datos para guardar los titulares y la fecha.

ayer=pd.read_csv("Publico.csv",sep=";",encoding='UTF-8',index_col="Indice")
hoy=ayer.append(Publico)
#Vamos fusionando los datos que ya teníamos con los recogidos hoy.

hoy.to_csv("Publico.csv",encoding='UTF-8',index_label="Indice")
#Guardamos en csv la fusión de lo acumulado y las nuevas noticias de hoy.

```

Código de la clasificación política de titulares:

```
import os
import pandas as pd
import numpy as np
import re
import nltk
from itertools import combinations
from nltk.corpus import stopwords

Partidos={"PP":["(PP)", "Rajoy", u"Santamaría", "Soraya", "Casado", "Hernando",
               "Cospedal", u"Maíllo", "Albiol", "Maroto", "Cifuentes", "Levy",
               "Garrido", u"Feijóo", u"García Egea", "Montserrat"],
          "PSOE":["(PSOE)", "PSC", u"Sánchez", u"Ábalos", "Pedro Duque", u"Díaz",
                 "Calvo", "Borrell", u"Patxi López", "Lastra",
                 "Montero", "Marlaska", u"Ábalos", u"Montón",
                 "Planas", "Robles", u"Celaá"],
          "Podemos":["Iglesias", "Irene Montero", "Echenique", u"Errejón",
                    "Carmena", "Colau", u"Garzón", "Espinár", u"Compromís",
                    "Comuns", "Bescansa"],
          "Ciudadanos":["(Cs)", "Rivera", "Arrimadas", u"Cantó", "Girauta",
                       "Villegas", u"Villacís", "Arcadi Espada"]}]

Medio=pd.read_csv("ElDiario.csv", sep=";", encoding='UTF-8')
#Cargamos el archivo que contiene los titulares.

Medio["Partido"]=np.repeat("None", len(Medio["Noticia"]))
#Creamos una nueva columna que contendrá el partido al que pertenece.

Medio["Partido Literal"]=np.repeat("None", len(Medio["Noticia"]))
#Si es noticia de referencia directa a un partido.

Medio1=Medio.loc[~Medio["Noticia"].duplicated(keep="first")]
#Nos quedamos solo con las noticias que no están duplicadas.

Valoresdic=(Partidos["PP"]+Partidos["PSOE"]+
            Partidos["Podemos"]+Partidos["Ciudadanos"])
#Conjunto de todos los valores del diccionario.

Valoresdic1=set(Valoresdic) | set(["PP", u"Génova", "PSOE", "Moncloa",
                                   "Podemos", "Ciudadanos", "Cs", "CS", "cs"])
#Conjunto de valores del diccionario y referencias literales a partidos.

#StopWords
ST=stopwords.words('spanish')

keywords={}
#Busqueda de palabras clave para el diccionario.
```

```

for n in Medio["Noticia"]:
#Llenamos keywords para su analisis. Aproximadamente 1 hora por medio.
    splitted=[x for x in re.split(';+|,| |-|:"|»|!+|¿|«|'|\[?]',n) if x!=""]
    #Multi separador.
    for w in splitted:
        if str.isupper(w[0:1].encode('utf-8')) and w not in ST:
            if w not in keywords.keys():
                keywords[w]=1
            else:
                keywords[w]+=1
        else:
            continue

print [x+": "+str(keywords[x]) for x in keywords.keys() if keywords[x]>3]
#Observamos si hay alguna palabra que incorporar a "Partidos".

def key(palabra):
#Funcion que devuelve la key de un valor del diccionario.
    for k,v in Partidos.iteritems():
        if palabra in [x for x in Partidos[k]]:
            return k

conectores={}
for i in Mediol["Noticia"]:
#Estudio de noticias que conectan distintos partidos.
    palabras=[x for x in re.split(';+|,| |-|:"|»|!+|¿|«|'|\[?]',i) if x!=""]
    interseccion=set(palabras) & set(Valoresdic)
    if len(interseccion)>1:
        posiciones=[]
        for x in interseccion:
            posiciones.append(palabras.index(x))
        for a,b in combinations(sorted(posiciones),2):
            dist=b-a
            if (dist==2) and (key(palabras[b])!=key(palabras[a])):
                print i,Mediol.loc[Mediol["Noticia"]==i]
                if palabras[b-1] in conectores.keys():
                    #Insertamos clave de diccionario y sumamos 1 en valor.
                    conectores[palabras[b-1]]+=1
                else:
                    conectores[palabras[b-1]]=1
            else:
                continue
    else:
        continue

```



```

afterkeyword={}
for n in Mediol["Noticia"]:
    """Vamos a buscar las palabras que suelen ir justo delante
    de las pertenecientes al diccionario."""
    try:
        splitted=[x for x in re.split(';+|,| |-|:|"+|»|!+|¿|«|[[?]',n)
                  if x!=""]
        for i in splitted:
            if i in Valoresdicl:
                if splitted[splitted.index(i)+1] in
                afterkeyword.keys():
                    afterkeyword[splitted[splitted.index(i)+1]]+=1
                    break
            else:
                afterkeyword[splitted[splitted.index(i)+1]]=1
                break
        elif (i in ["Gobierno","Ejecutivo"] and
              splitted[splitted.index(i)-1]=="El") or
        (i=="Partido" and splitted[splitted.index(i)+1] in
         ["Popular","Socialista"]) or
        (i==u"García" and splitted[splitted.index(i)+1]=="Egea") or
        (i=="Arcadi" and splitted[splitted.index(i)+1]=="Espada") or
        (i=="Patxi" and splitted[splitted.index(i)+1]==u"López"):
            if splitted[splitted.index(i)+2] in afterkeyword.keys():
                afterkeyword[splitted[splitted.index(i)+2]]+=1
                break
            else:
                afterkeyword[splitted[splitted.index(i)+2]]=1
                break
        else:
            continue
    except:
        continue

for z in afterkeyword.keys():
    if afterkeyword[z]>3:
        print z,afterkeyword[z]

liter=["pide","dice","afirma","anuncia","ve","considera","responde",
       "asegura","avisa","recalca","cree","acusa","rechaza","niega",
       "avisa","llama","replica","plantea","insta","critica","aclara",
       "reitera","reconoce","alerta","advierde","reprocha"]

```

Hemos estudiado las palabras que se sitúan justo delante de los políticos o partidos pertenecientes al diccionario. El objetivo es buscar las que hacen referencia a lo que manifiestan para tenerlo en cuenta en la clasificación siguiente.

```

Fila=0
for n in Mediol["Noticia"]:
    try:
        splitted=[x for x in re.split(';+|,| |-|"+|»|!+|¿|«|'|'?',n)
                    if x!=""]
        posicion=0
        for i in splitted:
            if (i in Partidos["PP"] and splitted[posicion+1] in liter)
            or (i in ["PP",u"Génova"] and splitted[posicion+1] in liter)
            or (i in ["Partido",u"García"] and splitted[posicion+1] in
                ["Popular", "Egea"] and splitted[posicion+2] in liter):
                #Titulares que hacen referencia a lo que manifiestan.
                Mediol.iat[Fila,3]="PPD"
                break
            elif (i in Partidos["PP"]) or (i in ["PP",u"Génova"]) or
            (i in ["Partido",u"García"] and splitted[posicion+1] in
                ["Popular", "Egea"]):
                Mediol.iat[Fila,3]="PP"
                break
            elif (i in Partidos["PSOE"] and splitted[posicion+1] in liter) or
            (i in ["PSOE", "Moncloa"] and splitted[posicion+1] in liter) or
            (i=="Partido" and splitted[posicion+1]=="Socialista" and
                splitted[posicion+2] in liter) or
            (i=="Patxi" and splitted[posicion+1]==u"López" and
                splitted[posicion+2] in liter) or
            (i=="Pedro" and splitted[posicion+1]=="Duque" and
                splitted[posicion+2] in liter) or (i=="Patxi" and
                splitted[posicion+1]==u"López"
                and splitted[posicion+2] in liter) or
            (i=="El" and splitted[posicion+1] in ["Gobierno", "Ejecutivo"] and
                splitted[posicion+2] in liter):
                #Titulares que hacen referencia a lo que manifiestan.
                Mediol.iat[Fila,3]="PSOED"
                break
            elif (i in Partidos["PSOE"]) or (i in ["PSOE", "Moncloa"]) or
            (i=="El" and splitted[posicion+1] in ["Gobierno", "Ejecutivo"]) or
            (i=="Partido" and splitted[posicion+1]=="Socialista") or
            (i=="Patxi" and splitted[posicion+1]==u"López") or
            (i=="Pedro" and splitted[posicion+1]=="Duque"):
                Mediol.iat[Fila,3]="PSOE"
                break
            elif (i in Partidos["Podemos"] and splitted[posicion+1] in liter)
            or (i=="Podemos" and splitted[posicion+1] in liter) or
            (i=="Irene" and splitted[posicion+1]=="Montero" and
                splitted[posicion+2] in liter):
                #Titulares que hacen referencia a lo que manifiestan.
                Mediol.iat[Fila,3]="PodemosD"
                break
            elif (i in Partidos["Podemos"]) or (i=="Podemos") or
            (i=="Irene" and splitted[posicion+1]=="Montero"):
                Mediol.iat[Fila,3]="Podemos"
                break

```

```

elif (i in Partidos["Ciudadanos"] and splitted[posicion+1] in liter)
or (i in ["Ciudadanos","Cs","CS","cs"] and
    splitted[posicion+1] in liter) or
(i=="Arcadi" and splitted[posicion+1]=="Espada" and
    splitted[posicion+1] in liter):
    #Titulares que hacen referencia a lo que manifiestan.
    Medioli.iat[Fila,3]="CSD"
    break
elif (i in Partidos["Ciudadanos"]) or (i in ["Ciudadanos","Cs","CS","cs"])
or (i=="Arcadi" and splitted[posicion+1]=="Espada"):
    Medioli.iat[Fila,3]="CS"
    break
elif (i[0:len(i)-1] in Partidos["PP"] and i[-1]==":") or
(i[0:len(i)-1] in ["PP",u"Génova"] and i[-1]==":") or
(i in ["Partido",u"García"] and
    splitted[posicion+1][0:len(splitted[posicion+1])-1] in
    ["Popular","Egea"] and splitted[posicion+1][-1]==":"):
    #Noticias que son citas textuales completas.
    ElDiariol.iat[Fila,3]="PPD"
    break
elif (i[0:len(i)-1] in Partidos["PSOE"] and i[-1]==":") or
(i[0:len(i)-1] in ["PSOE","Moncloa"] and i[-1]==":") or
(i=="El" and splitted[posicion+1][0:len(splitted[posicion+1])-1]
    in ["Gobierno","Ejecutivo"] and splitted[posicion+1][-1]==":") or
(i=="Patxi" and splitted[posicion+1][0:len(splitted[posicion+1])-1]==u"López"
    and splitted[posicion+1][-1]==":") or
(i=="Pedro" and splitted[posicion+1][0:len(splitted[posicion+1])-1]=="Duque" and
    splitted[posicion+1][-1]==":") or
(i=="Partido" and splitted[posicion+1][0:len(splitted[posicion+1])-1]=="Socialista"
    and splitted[posicion+1][-1]==":"):
    #Noticias que son citas textuales completas.
    ElDiariol.iat[Fila,3]="PSOED"
    break
elif (i[0:len(i)-1] in Partidos["Podemos"] and i[-1]==":") or
(i[0:len(i)-1]=="Podemos" and i[-1]==":") or
(i=="Irene" and splitted[posicion+1][0:len(splitted[posicion+1])-1]=="Montero"
    and splitted[posicion+1][-1]==":"):
    #Noticias que son citas textuales completas.
    ElDiariol.iat[Fila,3]="PodemosD"
    break
elif (i[0:len(i)-1] in Partidos["Ciudadanos"] and i[-1]==":") or
(i[0:len(i)-1] in ["Ciudadanos","Cs","CS","cs"] and i[-1]==":") or
(i=="Arcadi" and splitted[posicion+1][0:len(splitted[posicion+1])-1]=="Espada"
    and splitted[posicion+1][-1]==":"):
    #Noticias que son citas textuales completas.
    ElDiariol.iat[Fila,3]="CSD"
    break
else:
    posicion+=1
    continue
Fila+=1
except:
    Fila+=1
    continue

```

Mediante este enorme bucle “for” clasificamos los titulares según la primera aparición de una palabra correspondiente al diccionario de partidos creado.

```

Mediol.drop(Mediol[Mediol.Partido=="None"].index, inplace=True)
#Nos quedamos solo con los titulares referentes a los cuatro partidos analizados.

duplicada=[]
recicle=[]
for x in Mediol["Noticia"]:
    #Detectamos si hay alguna duplicada no exacta para finalizar.
    duplicada.append(x)
    for i in Mediol["Noticia"]:
        if (x!=i) and (i not in duplicada):
            frase=[n for n in re.split(';+|,| |-|:|"|»|!|+|¿|«|'|'?',x) if n!=""]
            splitted=[d for d in re.split(';+|,| |-|:|"|»|!|+|¿|«|'|'?',i) if d!=""]
            contador=0
            coincidencia=0
            for w,j in zip(frase,splitted):
                contador+=1
                if w==j:
                    coincidencia+=1
                elif contador==5:
                    break
            else:
                break
            if coincidencia>4:
                print x,i
                recicle.append(i)

for r in recicle:
    Mediol.drop(Mediol[Mediol.Noticia==r].index, inplace=True)

Mediol.to_csv("Probandoresultados.csv",encoding='UTF-8')
#Guardamos los resultados para su posterior análisis de sentimiento.

```

Cabe mencionar que todo el código referente a la clasificación de los titulares se ejecuta para cada archivo CSV correspondiente a los distintos medios de comunicación. Para ello, se debe cambiar el parámetro de la dirección del CSV.

Para terminar con el Apéndice vamos a mostrar el código referente al análisis de sentimiento de los titulares, si bien, omitimos la parte relacionada con los cálculos de los porcentajes, ya que no tienen más dificultad que localizar los datos en los distintos paneles de datos y realizar sumas o divisiones.

El código que mostramos a continuación debe ejecutarse para CSV que contiene los titulares de los distintos medios de comunicación.

Código del análisis de sentimiento:

```
# -*- coding: cp1252 -*-

import pandas as pd
import numpy as np
from meaningcloud import SentimentRequest, SentimentResponse

medio=pd.read_csv("Publicoclass.csv",encoding='utf-8',sep=',')
#Cargamos el archivo con los titulares clasificados en un panel.

medio["SentimientoAPI"]=np.repeat("None",len(medio["Noticia"]))
#Creamos la columna que va a contener el sentimiento.

headers={'content-type': 'application/x-www-form-urlencoded'}
#Formato de la respuesta.

APIkey="6d3ebe916158e5a693842a16a2a9ff9b"
#Key para usar la aplicacion.

url="https://api.meaningcloud.com/sentiment-2.1"
#Url de la aplicacion.

Fila=0
for x in medio["Noticia"]:
    #LLamamos a la aplicacion y guardamos resultado en el panel.
    try:
        frase=x
        request=request=SentimentRequest(key=APIkey,lang='es',
                                          txt=frase.encode('utf-8'),txtf='plain',
                                          url=url,extraheaders=headers).sendReq()

        #Llamada a la aplicacion.
        medio.iat[Fila,6]=SentimentResponse(response=request).getGlobalScoreTag()
        Fila+=1
    except:
        print "Existe un problema con: ",x
        Fila+=1
        continue

"""N=Negativo, NEU=Neutral, P=Positivo, P+=Fuertemente Positivo,
N+=Fuertemente Negativo, NONE= Sin Sentimiento"""

medio.to_csv("PublicoAPISENT.csv",encoding='UTF-8')
#Guardamos resultados en csv.
```