

Grado en Sistemas de Información



Trabajo Fin de Grado

Elaboración y análisis de información de una red social de la política española a través de la recolección de datos en Twitter



ESCUELA POLITECNICA
SUPERIOR

Autor: Daniel Ortega Expósito

Tutor/es: Salvador Sánchez Alonso

Cotutor/es: Marçal Mora Cantallops

Julio 2019

UNIVERSIDAD DE ALCALÁ
Escuela Politécnica Superior

GRADO EN SISTEMAS DE INFORMACIÓN

Trabajo Fin de Grado

Autor: Daniel Ortega Expósito

Tutor/es: Salvador Sánchez Alonso

Cotutor/es: Marçal Mora Cantallops

TRIBUNAL:

Presidente:

Vocal 1º:

Vocal 2º:

FECHA:

Agradecimientos

Comienzo agradeciendo a todas aquellas personas que han formado parte de mi vida durante estos cuatro años. A las que empezaron esta andadura conmigo y no la acabaron y a las que llegaron más tarde y aún siguen, pero sobre todo a aquellas que sé que no se irán. También a esa persona que durante años me lo dio todo.

A Salvador por introducirme en el mundo de las redes sociales, siempre te estaré agradecido. A Marçal por enseñarme que el correo de la universidad no es un chat ni una visita al psicólogo. En general a los dos, por permitirme aplicar lo conocido hasta ahora en un trabajo del que sentirme orgulloso y por su incondicional ayuda durante estos meses. Sin su apoyo, este trabajo no tendría sentido.

A todos aquellos profesores que han dedicado un minuto a nuestra formación con ilusión y profesionalidad. En especial a Esperanza, Luis e Iván.

A mis compañeros de clase Alejandro, María, Paula y Daniela, aunque dejaron de serlo hace tiempo para convertirse en amigos. Por hacer de cualquier momento de agobio una oportunidad para reírnos de la vida. Por pasar del amor al odio en cuestión de minutos. Por seguir siendo los reyes del mambo. Pensaba que nunca iba a ocurrir, pero ya echo de menos los trabajos en grupo. Que lo que unió la universidad, lo refuercen viajes como el de Lanzarote y no lo separe nadie jamás.

Hago especial mención a dos compañeros. Alejandro, por ser mi compañero incondicional no solo durante estos cuatro años, si no desde el instituto. Por ser el apoyo necesario para cualquier persona cuando vienen malas y la mejor compañía para disfrutar de las buenas. Aunque a veces abrumas a la gente con tu pesimismo, eres único y lo malo es que poca gente lo sabe. Rápido te has convertido en un amigo para toda la vida. Así que eso, no te vayas nunca.

A María, por convertirse este año en una persona muy especial en mi vida. No me dejaste caer cuando iba cuesta abajo y es algo que guardaré para siempre. Por ser mi compañera de TFG en la biblioteca y comida en la cafetería. Por pasar del agobio al pasotismo en la misma frase. Por ser la mejor confidente posible. No cambies nunca, eres un ser de luz. Nunca dejes que te apaguen.

A mis amigos fuera de aquí, imposible verme en esta situación sin tener cerca a gente con tanta calidad humana. Digan lo que digan, nos queda lo mejor.

Por último, a mi familia, ese apoyo que parece que siempre está y no se valora demasiado, pero cuando falta te das cuenta de todo lo que abarca. Gracias a mis padres y a mi hermano, os quiero.

Gracias a ti también, abuelo.

Resumen

Como refleja el título del trabajo, el proyecto consiste en la elaboración de la red política española a través de la recolección de información en Twitter, en este caso tweets. Esto consiste en la elección de los usuarios más relevantes del ámbito político en España y su posterior puesta a prueba ante eventos de carácter político durante la campaña electoral previa a las elecciones generales de 2019 en España. El proyecto tiene como objetivo la detección de los usuarios y movimientos más relevantes producidos durante estos eventos. Por último, se plantean las consecuencias políticas del contenido compartido en Twitter por parte de los integrantes de la red.

Abstract

As reflected in the title of the work, the project consists on the development of the Spanish political network through the collection of information on Twitter, in this case tweets. This consists of the election of the most relevant users of the political sphere in Spain and its subsequent testing against political events during the electoral campaign prior to the 2019 general elections in Spain. The project aims to detect the most relevant users and movements produced during these events. Finally, the political consequences of the content shared on Twitter by the members of the network are considered.

Palabras clave

Redes sociales, ARS, política, Gephi, Python, NetworkX.

Índice

Contenido

1. Introducción	1
2. Motivación	3
3. Estado de la cuestión	5
4. Objetivos	9
4.1. Objetivo general	9
4.2. Objetivos específicos:	9
5. Planificación	10
6. Metodología	12
6.1. Requisitos previos	12
6.2. Elaboración de la red política española	13
6.3. Recogida de datos en Twitter	14
6.4. Almacenamiento	15
6.5. Procesado de datos	15
6.6. Representación gráfica de la información	15
6.7. Análisis	16
7. Elaboración de la red	17
7.1. Requisitos previos	17
7.2. Proceso de elaboración	18
7.3. Criterios de elaboración	22
7.4. Tipología de usuarios	24
8. Recogida de datos en Twitter	27
8.1. Streaming API	28
8.2. API REST	33
8.3. Búsquedas complementarias	35
9. Almacenamiento	37
10. Procesado de datos	41
11. Construcción de grafos	44
12. Representación gráfica	47
13. Cálculo de la polaridad en la red	53
14. Análisis estático de la red	55
14.1. Estructura	55
14.2. Centralidad	57
14.3. Prestigio	61
14.4. Subredes	66

14.5.	Alternativas en la elaboración de la red	72
14.6.	Resumen	76
15.	Análisis dinámico de la red	78
15.1.	Debate electoral en RTVE el 22 de abril de 2019.....	78
15.2.	Debate electoral en Atresmedia el 23 de abril de 2019	85
15.3.	28 de abril - Día de elecciones.....	91
16.	Caso de estudio: Polaridad y tendencia de la red política española durante el período de campaña electoral	96
17.	Discusiones	107
18.	Conclusiones	110
19.	Bibliografía.....	112

1. Introducción

Actualmente nos hallamos en una sociedad en la que la tecnología ha normalizado la capacidad de estar conectado de diferentes formas con alguien o algo.

Se conoce a las redes sociales como aquellas aplicaciones o programas que utilizamos para interactuar con otros. Es un error que se comete muy a menudo, ya que a lo que de verdad nos estamos refiriendo es a las plataformas sociales.

De otro modo, una red social es, según Barnes, “Un conjunto de puntos, algunos de los cuales están unidos por líneas. Los puntos son personas o grupos y las líneas indican que los individuos interactúan mutuamente. Así, se podría pensar que el conjunto de la vida social genera una red de este tipo”. [1]

La tecnología no ha creado las redes sociales como tal, sino que ha configurado otro modo de relacionarse. Además, ha proporcionado el poder suficiente como para permitir el avance en la investigación en el campo de ARS (Análisis de redes sociales).

Es año de elecciones en España. Con una tensión generalizada por temas de diverso carácter, el ámbito político comienza a interesar a un considerable sector de la población.

Cada vez más, se ve como los partidos políticos hacen uso de cualquier herramienta para su campaña electoral. Es una campaña de marketing más, pero ahora no solo se limitan a enviar papeletas y su propuesta a casa del ciudadano. Es un hecho que el uso de las plataformas sociales tiene un gran impacto sobre la población. Año tras año, la tecnología entra a formar parte de nuestras vidas con más firmeza. Por ello, los partidos políticos no han pasado por alto esta evolución y aprovechan la estrecha relación entre ciudadano y tecnología para divulgar sus ideas.

Una de las plataformas sociales más seguidas por los adeptos a la política es Twitter, donde los usuarios pueden interactuar directamente con otros usuarios independientemente de la tipología de este. Sigue una metodología similar a la de los SMS, mensajes cortos a uno o más usuarios, solo que, en este caso, la información suele ser pública. Quizás sea esto lo que atrae tanto al ciudadano como al partido político en todas sus formas, ya que existe la cercanía suficiente para llevar a cabo tu cometido a la par que das la imagen que deseas.

En este trabajo vamos a presentar una imagen más ideográfica de la red política en España a través de la integración del análisis de redes sociales en la política. Para ello, se elaborará una red representativa de la política en España a través de la recolección de información en Twitter.

Se utilizarán herramientas de visionado de redes sociales como Gephi y el módulo Networkx de Python.

Son innumerables las herramientas que existen hoy en día para el análisis de redes sociales. Es cierto que la comunidad suele quejarse de las pocas actualizaciones en cuanto a plugins

se refiere y al poco rango de mejora que ha experimentado la herramienta en los últimos años, pero Gephi continúa siendo uno de los principales recursos de los interesados en el análisis de redes sociales a la hora de poner en práctica sus proyectos de investigación, independientemente del tamaño de este.

Por lo tanto, la intención de este proyecto es poder aportar, en materia de investigación y desarrollo, algún punto de vista nuevo sobre lo ya existente. Para ello, se define un caso de estudio tras lo llevado a cabo durante el trabajo.

2. Motivación

Desde una temprana edad me ha llamado la atención cuestionar cada opinión que escuchaba, fuese en voz alta o íntimamente; la segunda opción en la mayoría de las ocasiones.

Rápido me atraieron los debates y discusiones sobre política, puestos en práctica, comúnmente, con mi familia y amigos. Era de los escasos momentos donde mi retracción, propia de un chico tímido, desaparecía.

Poco más tarde llegaron las plataformas sociales como hoy las conocemos. Para un joven ‘millennial’, como usualmente se nos conoce, resulta una ‘commodity’ más. Como quien se levanta y apaga el despertador, hemos adquirido la costumbre de ver qué nuevas relaciones hemos establecido desde por la mañana.

Una de las plataformas sociales, o redes sociales cómo públicamente se las conoce, que más utilizo es Twitter. Una ocasión única para aquellos a los que nos apasiona litigar cada opinión e iniciar una contienda sana.

En 2015 comencé a estudiar el Grado en Sistemas de Información en Alcalá de Henares, en la Universidad de Alcalá. Como muchos otros compañeros de mi instituto, aún tenía dudas sobre qué carrera cursar. A diferencia de otros grados, vi en este, una oportunidad para proyectar todo lo aprendido en otros ámbitos de la vida. Creo que no me equivoqué demasiado.

La peculiaridad de este grado es que no es una ingeniería como el resto de las carreras de la facultad. Una buena parte de las asignaturas está relacionada con materias más allá de lo estrictamente necesario para considerarse ingeniero. La mayoría se acercan a la economía y la gestión de equipos. Creo que te llega a abrir más puertas que el resto de los grados debido a su variedad de contenidos.

Esta singularidad me ha permitido no cerrarme en ningún campo de la carrera y cuestionar hasta el último momento por donde continuar mi carrera profesional. Gracias a esto me he planteado varios proyectos de fin de carrera.

Pasados los años y las asignaturas, me llegó la hora de decidir que trabajo de fin de grado quería hacer. No pretendía hacer algo de lo que no me sintiese orgulloso y que no aportase algo de valor a quien lo leyese.

En un primer momento me interesé por los trabajos ofertados en materia de *machine learning*, pero al ver que se ofrecían temas relacionados con la extracción y análisis de datos en plataformas sociales y que además se aplicaba al ámbito de la política, no dudé en contactar con los tutores responsables.

El hecho de tener en pocas semanas unas elecciones generales, municipales y europeas, hizo que me pareciese una oportunidad única para proyectar todo lo aprendido en estos cuatro

años en un trabajo que de verdad me sedujese y motivase a dejar mi grano de arena en el campo del ARS (Análisis de Redes Sociales).

En definitiva, a continuación, se podrá leer mi trabajo. Espero que sea del agrado del lector y le aporte aquello que anda buscando.

3. Estado de la cuestión

En ‘Models and methods in social network analysis’ [2], se define una red social como “un grupo de personas que de forma agrupada o individual se relacionan entre sí y, a su vez, con otras personas de su medio. Las relaciones entre ellos pueden ser muy diversas y pueden existir muchas o pocas personas en cada red”. Se puede sostener entonces, que cualquier interacción como consecuencia de la relación entre dos individuos es objeto de análisis.

Para Barnes [1], “en una red social participan principalmente dos aspectos, los nodos o vértices que simbolizan a las personas que conforman la red social y las aristas o enlaces que representan las relaciones entre estos”. Es un hecho, por lo tanto, que debe existir un término que abarque este estudio.

Según L. Freeman, “En las ciencias sociales, el enfoque estructural que se basa en el estudio de la interacción entre actores sociales se llama análisis de redes sociales.” [7].

La representación de estas redes aplica una serie de pautas siguiendo la Teoría de Grafos. La teoría de los grafos (Ore, 1195 [3] y Biggs, 1974 [4]) es la rama de la Matemática Discreta (Bujalance y otros, 1993) [5] que se ocupa de este tipo de problemas. La conectividad entre los elementos de un conjunto es el objetivo fundamental de la teoría de grafos.

La tecnología es capaz de proporcionar la capacidad de poder analizar estas relaciones, hecho que antes parecía irrealizable manualmente. De hecho, a lo largo de los años, numerosos estudios e investigaciones han quedado estancados por la incapacidad computacional para corroborar sus resultados. Por ejemplo, en la década de los 50, Ithiel de Sola Pool (MIT) y Manfred Kochen (IBM) se propusieron demostrar la teoría de los seis grados matemáticamente. Esta teoría asegura que es posible acceder a cualquier persona del planeta en tan solo seis saltos. Aunque eran capaces de enunciar la cuestión "dado un conjunto de N personas, ¿cuál es la probabilidad de que cada miembro de estos N esté conectado con otro miembro vía $k_1, k_2, k_3, \dots, k_n$ enlaces?", después de más de dos décadas todavía eran incapaces de resolver el problema satisfactoriamente [6].

Esta falta de recursos, a lo largo de la historia, nos ha privado de satisfacer nuestro anhelo por conocer y descubrir nuevos hallazgos. A menudo, esto hace que no seamos conscientes de la inmensa cantidad de información que acaba escapándose en situaciones cotidianas, por muy breves que sean. Esa aspiración por captar determinada información hace que la técnica de análisis de redes sociales esté adquiriendo un impacto importante en diferentes ámbitos de la vida real.

Las redes sociales son susceptibles al análisis puesto que generan un comportamiento social. Este análisis es empleado en cuantiosos ámbitos cotidianos para el alcance de conclusiones que aporten valor al analista. El auge del desarrollo de la tecnología permite, cada vez más, ver como el análisis de redes sociales se aplica a diversidad de campos de estudio.

Podemos ver como cada vez existen más ejemplos de que el ARS está cogiendo forma en diferentes campos científicos:

- Andre, R. Pinheiro [8], en ‘Social Network Analysis in Telecommunications’ nos presenta una mirada al uso efectivo de redes sociales dentro de la industria de las telecomunicaciones para impulsar las relaciones con los clientes.
- Clemente, Filipe Manuel, F. M. Lourenço Martins, y R. S. Mendes [9], en ‘Social Network Analysis Applied to Team Sports Analysis’ nos explican cómo la teoría de grafos y el análisis de redes sociales se pueden aplicar al análisis de deportes de equipo. Tiene como objetivo proporcionar al lector una metodología sobre medidas de análisis de redes sociales para la extracción de información y aplicación, a través de análisis de interpretación, en competición oficial.
- Tayebi, Mohammad A., y Uwe. Glässer [10], en ‘Social Network Analysis in Predictive Policing: Concepts, Models and Methods’ se centran en las aplicaciones del análisis de redes sociales en la vigilancia predictiva. Para estos autores “la ciencia de los datos se usa para identificar la actividad criminal potencial al analizar las relaciones entre los delincuentes para comprender completamente los patrones de colaboración criminal”.

Estos son solo algunos ejemplos de aplicación del análisis de redes sociales en distintos ámbitos de nuestro mundo. Pero, cómo el número de relaciones que se crean a diario, existen infinidad de casos prácticos donde el ARS tiene cabida.

En [7] se referencia a una de las obras de L. Freeman (Encyclopedia of Psychology, 2000) donde ya dijo lo que venimos enunciando, dónde cito textualmente, “el campo de la red ha desarrollado importantes aplicaciones en la investigación sobre el estudio de la movilidad profesional, el impacto de la urbanización sobre los individuos, el mundo político y sistema económico, toma de decisiones comunitaria, apoyo social, resolución de problemas grupales, difusión e interconexión corporativa, sistemas de creencias, cognición social, mercados, sociología de la ciencia, intercambio y poder, consenso e influencia social, estudios de primates, comunicación mediada por computadora, estudios intra e interorganizacionales o campos como la salud donde se han estudiado enfermedades, particularmente el sida”.

Es evidente el amplio espectro del que dispone esta disciplina y de lo que aún es capaz de abarcar.

Una de estas áreas que menciona Freeman y que está alcanzando un rol más determinante es el mundo político. Prácticamente todos los partidos políticos actuales ya tienen en cuenta el uso de las plataformas sociales y las utilizan como una herramienta electoral más. Es un medio perfecto para establecer contacto con aquellos simpatizantes o potenciales votantes. Junto con el auge del desarrollo tecnológico, se presentan amplias oportunidades de investigación en el ámbito político.

Podemos encontrar varias referencias en torno a este campo como, por ejemplo, en ‘Explaining How Political Actors Gain Strategic Positions: Predictors of Centrality in State Reading Policy Issue Networks. Educational Policy’ [12], se explica cómo los individuos adquieren posiciones centrales en redes temáticas. En este caso los individuos son políticos de California, Connecticut, Michigan y Utah. En este estudio, los hallazgos sugieren que los actores estaban formando relaciones con otros actores para obtener acceso a organizaciones influyentes que podrían tener más recursos que los propios actores políticos.

También existen otras publicaciones que nos acercan aún más al objetivo de nuestro trabajo. En ‘A Longitudinal Social Network Analysis of German Politicians Twitter Accounts’[13] nos encontramos en una situación similar a la planteada en este trabajo. Los políticos se postulan para el cargo de la elección del Bundestag alemán. Este estudio se realiza durante el 2009, año de las elecciones. En el estudio se tomaron instantáneas diarias de la red en Twitter. Esta red estaba formada por más de 650 políticos en un plazo de 4 semanas. Rastrearón casi 10,000 conexiones de más de 650 cuentas de twitter y monitorizaron más de 240,000 tweets.

Usando métodos estadísticos de análisis de redes sociales sacaron algunas conclusiones interesantes:

- Existe una indicación de cierre entre los miembros de su propio partido
- La mayoría de las conexiones se establecen entre miembros del mismo partido mientras que las conexiones entre diferentes partidos están significativamente menos representadas,

Otros casos, como el planteado en [15], nos muestra la importancia del contenido de la información que transcurre por la red y su relación con el entorno en sucesos como el enjuiciamiento del presidente de Brasil.

En España también ha habido casos de estudio en relación con el planteamiento de nuestro trabajo. En ‘La campaña virtual en twitter: análisis de las cuentas de Rajoy y de Rubalcaba en las elecciones generales de 2011’ [14].

Durante la lectura del artículo de investigación, podemos encontrar una serie de referencias que nos ayudarán a entender mejor el contexto de la política española en Twitter. De acuerdo con el estudio realizado en Alemania, concluyen con “En este sentido, explotaron las posibilidades de la red de microblogging para interactuar, sobre todo, con ciudadanos. Menor fue la interacción con otros políticos, mayoritariamente del propio partido.”

Sin embargo, por otro lado, aporta algo más de información con “interacción muy escasa con periodistas o medios de comunicación, lo que muestra que Twitter no se empleó durante la campaña como vía de diálogo con la prensa”.

Hace casi 6 años de este estudio, con el paso del tiempo y la exponente aceptación de la tecnología en la sociedad, ¿podemos seguir afirmando este hecho?

Pero hay más, cito lo siguiente, “Rajoy utilizó bastante el retweet para difundir mensajes procedentes de cuentas de periodistas y de medios de comunicación, lo que matiza parcialmente el resultado mostrado anteriormente acerca de las referencias a la prensa por parte de los candidatos”. Se menciona que Rajoy interactúa con medios de comunicación, pero ¿cuáles y por qué?

Hasta ahora, se ha visto como se plantean casos de estudio sobre la política en redes sociales tanto en España como en el resto del mundo. Por otro lado, actualmente, en España se vive una situación excepcional; por primera vez en la historia política española, cinco partidos políticos se dividen casi la totalidad de la representación en las futuras elecciones. Es por ello por lo que, cualquier mala decisión tomada por un partido puede ser determinante para su futuro. Uno de estos factores clave es el uso de las redes sociales y la importancia de su contenido. Por ello, en este trabajo se plantea el análisis de sentimiento del contenido que los partidos políticos comparten a través de Twitter y sus posibles consecuencias de cara a las elecciones generales de 2019.

4. Objetivos

Para la consecución de este trabajo de fin de grado se puede definir un objetivo general del mismo y otros subobjetivos más específicos.

4.1. Objetivo general

El objetivo general de este trabajo consiste en elaborar una red completa de la política española a través de la comparación de actores únicos en otras plataformas sociales y la monitorización de sus cuentas en Twitter, obteniendo así, información con un alto grado de valor que permita proyectar una imagen más representativa de los principales actores del ámbito político en España.

4.2. Objetivos específicos:

- Elaborar una red inicial de la política española, identificando quiénes son los actores más relevantes en la misma y encontrando las relaciones entre los mismos.
- Interpretar estructura y comportamiento de la red ante dos casos de análisis de distinta naturaleza. La tipología de los casos difiere en el dinamismo de los datos obtenidos de la red. Se toman como referencia dos casos para mejorar el contraste de información y obtención de conclusiones.
 - o El primer caso de análisis servirá como puesta de relieve de la estructura de la red y la detección de los actores más importantes a priori. Únicamente se tomarán datos a partir de la conformación inicial de la red, es decir, una visión estática de esta.
 - o El segundo caso proyectará una imagen diferente a partir de la conducta de la red ante una serie de eventos de carácter político durante la campaña electoral.
- Plantear caso de estudio que ofrecerá el sentimiento de los tweets de cada usuario y sus consecuencias, facilitando así poder encontrar sucesos de interés o similitudes con los casos de análisis anteriores.

5. Planificación

A lo largo de la elaboración del TFG se van a llevar a cabo una serie de actividades para la consecución de este.

A continuación, se describen cada una de las actividades en cuestión, de modo que quede concretada la estructura del trabajo:

- Documentación: la fase de documentación abarcará prácticamente todo el proyecto. Desde su inicio hasta pocos días antes de su entrega. Se estima una duración aproximada de 16 semanas.
- Descarga y formación sobre las herramientas: esta fase consiste en la adquisición y configuración del entorno y herramientas necesarias para la consecución del trabajo. Estas herramientas son las siguientes:
 - o Gephi (versión 9.2)
 - o Python (versión 3.7)
 - o Librerías para Python como por ejemplo tweepy o pymongo
 - o MongoDB (versión 4.0.7) con MongoDB Compass incluida para la visualización de datos.

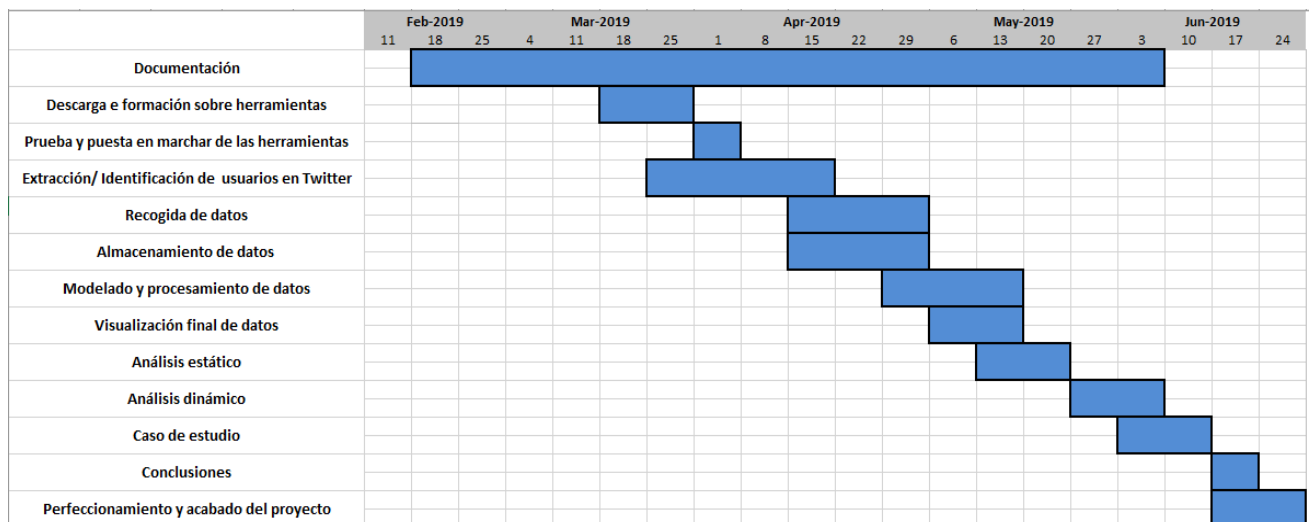
Se estima una duración aproximada de 2 semanas.

- Prueba y puesta en marcha de las herramientas: nos cercioramos de que las herramientas instaladas desempeñan su función de acuerdo con nuestras expectativas. Se estima una duración aproximada de 1 semana.
- Extracción/identificación de usuarios en la red Twitter: elección y recogida de usuarios relacionados con la red política española en Twitter en base a una serie de criterios explicados más adelante. Se estima una duración aproximada de 4 semanas. La razón de su largo transcurso reside en su especial laboriosidad durante el proceso de selección de usuarios.
- Recogida de datos: recolección de tweets a través de Python y la API de Twitter en momentos clave como, por ejemplo, los debates o el día de las elecciones. Se estima una duración aproximada de 3 semanas. Hay que tener en cuenta que con API rest se es capaz de extraer determinada información hasta siete días después de su creación. Así se puede explicar que, aproximadamente, desde el día de los debates hasta siete días después del día de las elecciones, exista un rango de tres semanas.
- Almacenamiento de datos: tras recoger los datos, necesitamos guardarlos en alguna base de datos, en este caso MongoDB. Es un proceso en paralelo con la recogida de datos, es decir, según llega la información es almacenada en la base de datos. Por lo tanto, al igual que la recogida de datos, se estima una duración aproximada de 3 semanas.

- Modelado y procesamiento de los datos: No nos interesa toda la información recogida. Necesitamos procesar los datos de tal forma que obtengamos aquella información con la que queremos trabajar. Se estima una duración aproximada de 3 semanas.
- Visualización final de la información: es la construcción de grafos a partir de la información relevante para poder iniciar el análisis determinado. Se estima una duración aproximada de 3 semanas.
- Análisis estático: análisis pormenorizado de la red política española elaborada con antelación. Se estima una duración aproximada de 2 semanas.
- Análisis dinámico: análisis de la red a través de la visualización en grafos construidos a partir de información recogida en momentos clave del proceso electoral. Se estima una duración aproximada de 2 semanas.
- Caso de estudio: puesta en práctica un caso de estudio a partir de los resultados obtenidos durante las fases previas del trabajo. Se estima una duración aproximada de 2 semanas.
- Conclusiones: elaboración de conclusiones tras los resultados de anteriores apartados. Se estima una duración aproximada de 1 semana.
- Corrección de errores y acabado del trabajo: Se estima una duración aproximada de 1 semana.

A modo de resumen, se esquematiza la planificación con el siguiente diagrama:

Figura 1. Planificación del trabajo



Fuente: Elaboración propia

6. Metodología

Como se ha mencionado en el apartado de planificación, el trabajo consta de una serie de actividades por las que se ha de pasar para la consecución del objetivo y, por tanto, del proyecto. A continuación, se procede a su explicación con mayor detalle.

6.1. Requisitos previos

En primer lugar, para comenzar con el trabajo, necesitó cumplir con una serie de requisitos. Para ello:

- Se creó una cuenta de desarrollador en Twitter. Esto fue necesario para tener acceso de uso a la API. Una vez se tuvo la cuenta, se registró una nueva aplicación, con la que nos proporcionaron nuestras credenciales y el permiso para utilizar las diferentes funcionalidades de desarrollador en Twitter.
- Se utilizó la aplicación Netvizz de Facebook, una herramienta que ayuda a analizar diferentes secciones de la plataforma de Facebook, principalmente páginas, con fines de investigación. Para utilizarla, antes se instaló la aplicación desde la plataforma, por lo que también se necesitó crear una cuenta en Facebook.
- En tercer lugar, se instaló MongoDB. Esta herramienta fue útil para el almacenamiento de datos obtenidos a partir de Twitter. A partir de ella gestionaremos las importaciones a otra herramienta de modelado y visualización de datos.
- Esta última herramienta mencionada es Python, con la que no sólo se pudo generar grafos a partir de datos, sino que fue la herramienta que nos ofreció la capacidad de extraer información de Twitter. Los grafos generados fueron visualizados y tratados en otra herramienta.
- Esta herramienta es Gephi. Es open-source y está desarrollada en Java. Nos permite visualizar y analizar grandes gráficos de red. Usa un motor de renderizado 3D para mostrar gráficos en tiempo real y permite explorar, analizar, filtrar, clusterizar, manipular y exportar diversos tipos de gráficos. Con Gephi se ha visualizado la red política española de una forma clara y completa, lo que ha facilitado la intención de analizar dicho mapa y sacar alguna conclusión sobre el comportamiento de los partidos políticos en todas sus formas (partido como marca, como conjunto de políticos, representación del partido por un determinado político, roles de políticos en plataformas sociales, relación de lo anterior con actores externos, etc.).

Una vez se tuvo a nuestra disposición las herramientas anteriores, dio comienzo el proyecto.

6.2. Elaboración de la red política española

Se inició el trabajo con el objetivo de elaborar la red política española. Para ello se necesitó desarrollar una serie de tareas.

La primera tarea consistió en la obtención de la red política española en Facebook. El proceso fue llevado a cabo por partes, es decir, el mismo cometido por cada partido político de los cinco más influyentes en España. Estos partidos son: Partido Socialista Obrero Español (PSOE), Partido Popular (PP), Ciudadanos (Cs), Unidas Podemos (UP) y VOX.

Para ello, se utilizó la aplicación Netvizz de Facebook. Esta aplicación fue útil para una primera toma de contacto y puesta en marcha del proceso de elaboración de la red política española.

Para tener permiso de uso de Netvizz, primero se necesitó crear una cuenta en la propia plataforma social. Se utilizó esta herramienta porque es una forma fácil y precisa de obtener el entorno de influencia de cada partido político. Netvizz cuenta con una serie de funcionalidades o módulos. En este trabajo se ha limitado a utilizar ‘page like network’. Este módulo de Netvizz devolvió una visualización en forma de grafo de los actores que intervienen en el entorno del usuario en cuestión. En este caso, esos usuarios son los partidos políticos que se han establecido. El orden en el proceso de extracción de datos ha sido el siguiente: UP, VOX, PSOE, PP y Cs. Este orden se ha establecido según la densidad, a priori, de usuarios relevantes en plataformas sociales, de mayor a menor. Hay que tener en cuenta que la red de cada partido no se limita a usuarios que practiquen la política habitualmente, sino que pueden ser otro tipo de usuarios como periodistas, medios de comunicación o ‘influencers’.

Para indicar de qué usuario queremos obtener la red, se obtuvo su número identificativo en Facebook a través de una página facilitada por Netvizz. Por último, antes de solicitar dicho grafo, se tuvo que definir qué profundidad se quería establecer.

El resultado obtenido fue visualizado en Gephi, dónde, en su apartado de ‘laboratorio de datos’ se pudo exportar la tabla con los datos de la red subred resultante por cada partido político. La exportación se realizó en formato .csv. Este archivo fue importado en una hoja de cálculo, donde se inició el proceso de detección de usuarios en Twitter, descarte de usuarios no relevantes y agregación de aquellos que sí. Ese filtro se realizó en función de unos criterios que serán explicados en el apartado de desarrollo y no en el de metodología.

En paralelo, se llevó a cabo otro procedimiento para la agregación de nuevos usuarios y así poder perfeccionar aún más la red. Para dicho proceso no se siguió el mismo método de obtención de datos, sino que se recogió, a través de un script en Python y haciendo uso de la librería tweepy, a los representantes de cada partido para las elecciones generales de abril de 2019. Esta información es pública en Twitter y su autor es el usuario @dalvarez37, analista y consultor en ‘Social Media’.

Durante el desempeño de la tarea se almacenaron los nombres de usuario y enlace directo al perfil de cada actor de la red. Las hojas de cálculo están formadas por varios apartados en función de la tipología de los usuarios. Esto ha sido útil a la hora de detectar conclusiones que aporten valor al caso de estudio más adelante.

Una vez se detectaron y recogieron los usuarios más relevantes de cada uno de los partidos políticos más influyentes, se dio paso al mismo proceso, pero con aquellos que, a priori, poseían un papel más secundario. Estos partidos políticos son los siguientes (de mayor a menor, por número de actores recogidos): ERC (Esquerra Republicana de Catalunya), JxCAT (Junts per Catalunya), EHBildu (Euskal Herria Bildu), CCanaria (Coalición Canaria), Compromís, EAJPNV (Partido Nacionalista Vasco) y EnMarea.

Para estos partidos políticos solo se utilizó el segundo método de obtención de datos debido, en primera instancia, a su menor representación en la red global de la política española.

Como cuando se ha utilizado Netvizz, en este caso también se ha realizado un perfeccionamiento del resultado obtenido, añadiendo usuarios que no hayan sido obtenidos y el descarte de aquellos que no aporten valor.

Una vez se dispuso de la red, aunque susceptible a pequeños cambios, comenzó la siguiente fase.

6.3. Recogida de datos en Twitter

En esta fase el objetivo ha sido extraer información de Twitter en relación con la red que se ha determinado con anterioridad, en concreto aquellos tweets en los que han intervenido alguno de los integrantes de la red. Esta información ha sido útil para entender mejor el contexto de la red y conocer cómo se comporta ante eventos de carácter político como, por ejemplo, un debate o el día de las elecciones. Esta información se ha recogido en el transcurso de los debates de RTVE y Atresmedia realizados durante la campaña electoral previa a las elecciones generales de 2019, además del día de las elecciones.

Durante esta fase se ha utilizado Python como lenguaje para la extracción y MongoDB como herramienta de almacenamiento de esta. Sin embargo, este proceso se ha dividido en dos subprocesos según la tipología en la extracción de datos.

El primer método ha consistido en utilizar el Streaming API, con el que se han obtenido tweets en tiempo real mientras que, con API Rest, se han recogido tweets con una antigüedad máxima de siete días.

Es importante recalcar que, para el correcto funcionamiento de los scripts, se ha necesitado instalar e importar ciertas librerías. Algunos ejemplos de librerías son tweepy (uso de la API de Twitter en Python) o pymongo (uso de MongoDB en Python).

6.4. Almacenamiento

En paralelo a la fase de recogida de datos en Twitter, también se ha llevado a cabo la de almacenamiento de tweets durante los debates de RTVE y Atresmedia y el día de las elecciones.

En el propio script se ha definido como se han de guardar los tweets. Así mismo, se ha definido en qué base de datos y colección de MongoDB se quiere guardar la información recolectada.

Es un proceso muy significativo, ya que MongoDB ha servido cómo repositorio de toda la información recogida y se ha tenido la oportunidad de tratarla cómo y cuándo se quería.

6.5. Procesado de datos

El siguiente objetivo fue representar gráficamente la información. Para ello se utilizó el módulo de Python, Networkx. Pero, antes de nada, para que el resultado fuese fiable, como se precisaba de información relevante, se realizó un procesado de toda la información recogida.

Para la creación del grafo inicial se esbozó la red de un modo estático, es decir, sin tener en cuenta comportamientos momentáneos de la red como, por ejemplo, los producidos durante un debate. Para ello se necesitaron los ‘follows’ de cada usuario de nuestra red. Esto quiere decir a quién sigue cada participante de la red. Una vez se obtuvo, hubo que quedarse solo con aquellos ‘follows’ que perteneciesen a la red. Por lo tanto, se proyectó gráficamente como se relacionan los actores de nuestra red en base a quien sigue a quien en Twitter. Esto se puede conseguir haciendo uso de la API de Twitter. Aunque la API tiene una serie de restricciones fijas (15 llamadas por cada 15 minutos), en cuestión de unos días se puede tener almacenada dicha información.

Por otro lado, para el grafo representativo del caso de estudio se ha necesitado recoger aquellas relaciones interesantes de toda la información almacenada en la base de datos. Debía ser información valiosa y que aportase fundamentos suficientes como para sustentar posibles conclusiones.

6.6. Representación gráfica de la información

Esta fase consistió en la construcción de un grafo a partir de la información que se consideró relevante. Se utilizó el módulo Networkx de Python para la constitución de la red y los datos almacenados en MongoDB.

Aunque con Python también se puede descubrir y conocer algunos rasgos de la red creada, con un archivo graphml a partir del script en Python se puede visualizar la misma red en Gephi de una forma mucho más intuitiva. Por este motivo se decidió prescindir en la fase de visualización de NetworkX y optar por Gephi.

6.7. Análisis

En esta fase el objetivo fue analizar los datos recogidos a partir de la construcción de grafos. Se dividió en dos partes. En la primera se trató de contextualizar la red política española obtenida al comienzo del trabajo. Para ello, se dividió el análisis en dos partes:

- Análisis a priori de los datos recogidos en formato tabla para mejorar la comprensión de la estructura de la red.
- Visualizar (esbozo a partir de la información recogida) la red a través de Networkx y Gephi y su posterior análisis haciendo uso de las técnicas de análisis de redes sociales.

En la segunda se volvió a utilizar Gephi y Networkx para el análisis de nuestra red en eventos trascendentales y de carácter político mediante técnicas de análisis de redes sociales. En el grafo se determinó el carácter de estudio de este y dio pie a una puesta en marcha del caso de estudio, dónde se plasmó una idea a partir de todo lo cosechado con anterioridad.

7. Elaboración de la red

7.1. Requisitos previos

Para el desarrollo de la fase de elaboración de la red política española, se ha necesitado satisfacer una serie de requisitos.

Una vez se tuviese activada la cuenta de desarrollador tras la pertinente verificación de esta, se debía crear una aplicación. Esta aplicación es el producto del desarrollador que va a hacer uso de los datos de Twitter. Por ello, se debía expresar el motivo de esta y el modo en que iban a ser utilizados los datos proporcionados por Twitter.

Normalmente, Twitter no suele tardar más de una hora en validar la aplicación y dar acceso a la API a través de unas claves de autenticación, de las que se han hecho uso a lo largo del trabajo. Aunque existen casos en los que, comúnmente por falta de información o entendimiento de los datos cumplimentados en el formulario, se tarda algo más. De todas formas, en ese caso suelen volver a solicitar la información en cuestión.

También se ha utilizado la aplicación Netvizz de Facebook, una herramienta que ayuda a analizar diferentes secciones de la plataforma de Facebook, principalmente páginas, con fines de investigación. Para utilizarla, antes se debió instalar la aplicación desde la plataforma, por lo que también se necesitó crear una cuenta en Facebook.

Una vez se tuvo a disposición las herramientas comentadas en el apartado 6, se inició el proceso de elaboración de la red.

7.2. Proceso de elaboración

Como se ha descrito en el apartado 5.2, el primer paso para iniciar el trabajo es la elaboración de la red política española. Para ello se ha necesitado desarrollar una serie de tareas.

Se puede dividir el proceso completo en:

Extracción de subredes políticas a través de Netvizz: el objetivo de esta fase ha sido obtener el entorno de influencia de cada partido político. El módulo Netvizz devolvió una visualización en forma de grafo de los actores que intervenían en el ámbito del usuario en cuestión. En este caso, este usuario fue cada partido político que se definió.

Para acceder a Netvizz se debió buscar en el apartado de aplicaciones.

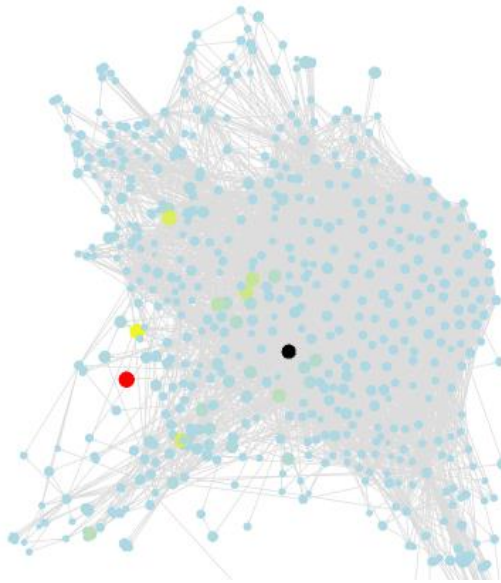
Una vez dentro de la aplicación, se hizo uso del módulo ‘page like network’.

Antes de ejecutar la búsqueda, se obtuvo el id del usuario que se quería analizar y la profundidad de la red, es decir, cuantos usuarios relacionados consecutivamente se querían en la red resultante.

Para averiguar el id (número identificativo en Facebook), la propia plataforma te proporciona un enlace. Tan solo hay que ingresar la dirección de la página en Facebook del usuario y te devuelve su id. La web se llama ‘lookup-id’.

En un primer momento, se pensó en fijar una profundidad de uno, pero el resultado no era muy representativo. Sin embargo, con una profundidad de dos, en ciertos casos, el resultado era excesivo. Finalmente, se prefirió optar por la segunda opción, aunque el trabajo de descarte de usuarios no relevantes fuese más laborioso.

Figura 2. Resultado obtenido tras la búsqueda de la página de Podemos con el módulo ‘page like network’



Fuente: Elaboración propia

La red resultante consta de 564 páginas en Facebook. Para importarlo a Gephi, tan solo se tuvo que descargar en formato zip, tal y como lo ofrece Facebook, y abrir el archivo GDF que contiene con Gephi.

Desde Gephi podemos visualizar la información de cada nodo de la red en el apartado ‘Laboratorio de Datos’. Desde este apartado se exportó un archivo csv con toda la información de la red.

Figura 3. Visualización de datos desde la ventana ‘Laboratorio de Datos’ en Gephi

Id	Label	username	category	post_activity	fan_count	talking_about_count
212534002288872	Podemos México	podemosmex	Community	0.0	2259	1
522931101151207	Podemos Motilla del Palancar	PodemosMotilladelPalancar	Community	0.0	415	1
242357829267813	Podemos Murcia Ciudad	podemosmurciaciudad	Community	0.08	948	115
1507163459504915	Circulo Salud Podemos Murcia	podemosmurciasanidadyssss	Community	0.0	918	0
482696115168680	Podemos Internacional	podemosnoticiasinternacional	Community	0.0	2817	1
664321210284774	Podemos Pozuelo	podemospozuelo	Community	0.01	847	0
506226472832954	Podemos Psicología	podemospsicologia	Community	0.0	7808	1
716522188369185	Podemos Región Leonesa	podemosregionleonesa	Community	0.0	895	3
570744529676994	Podemos Salamanca	PodemosSala	Community	0.1	4572	15
1438518336381730	Podemos Teruel	PodemosTeruel	Community	0.05	1800	15

Fuente: Elaboración propia

Una vez se dispuso del archivo csv con la red obtenida en Netvizz, se importó dicho archivo a una hoja de cálculo. Se ha elegido Excel para el almacenamiento de los usuarios de forma manual debido al conocimiento de la herramienta y su capacidad para exportar e importar la información deseada a otras plataformas. Es importante incidir en el riesgo del formato de la información, que puede variar según la herramienta que se utilice y que puede llevar a errores puntuales durante el proceso.

En la Figura 4 se aprecia la estructura del documento.

Figura 4. Estructura de los datos recogidos en Excel a partir de Netvizz

Id Facebook	Alias	Nombre de usuario	Categoría	Publicaciones/hora	Nº seguidores	Métrica de atención	URL
10513336322	The Guardian	theguardian	Media/News Company	255	8124542	816179	https://www.facebook.com/theguardian/
98444249426	Antena 3 Noticias	antena3noticias	News & Media Website	205	1276742	321235	https://www.facebook.com/antena3noticias/

Fuente: Elaboración propia

El siguiente paso fue crear una hoja nueva dentro del mismo documento dónde se exportaron los campos ‘Nombre de usuario’ y ‘URL’. Hecho esto, se inició la siguiente fase.

Búsqueda de usuarios homólogos en Twitter: Una vez se tuvieron los nombres de usuario y su URL a la página en Facebook, se inició la búsqueda manual de cada uno de los usuarios en Twitter. Los datos recogidos fueron los mismos: nombre de usuario y URL en Twitter.

En la Figura 5 se puede observar la estructura de la segunda hoja.

Figura 5. Estructura de la hoja de cálculo Excel con usuarios en Facebook y su homólogo en Twitter

FACEBOOK		TWITTER	
Nombre de usuario	URL	Nombre de usuario	URL
theguardian	https://www.facebook.com/theguardian/	guardian	https://twitter.com/guardian
antena3noticias	https://www.facebook.com/antena3noticias/	A3Noticias	https://twitter.com/A3Noticias

Fuente: Elaboración propia

Descarte de usuarios del total recogido: este proceso se ha realizado tanto durante la fase anterior como después de la misma. Se han descartado aquellas cuentas de Facebook que no tuviesen su equivalente en Twitter.

Después, también hubo que realizar un filtro de relevancia en la red, descartando de esta forma a aquellos que no fuesen representativos en la misma.

Ampliación y mejora de la red con nuevos usuarios relevantes: Por último, habría que saber detectar aquellos usuarios relevantes en Twitter que aún no se han agregado a la red, para así incorporarlos a esta.

Esta fase se ha llevado a cabo de dos modos diferentes:

1. De forma manual en Twitter, detectando conexiones con otros usuarios ya existentes en nuestra red. Lo ideal es empezar por la cuenta oficial del partido en cuestión e ir navegando a través de las sugerencias que te ofrece Twitter, que suelen ser cuentas directamente relacionadas con esta.

Por ejemplo, en la Figura 6, desde la cuenta oficial de Ciudadanos, podemos observar cómo sugiere seguir a usuarios muy cercanos al ámbito del partido como, por ejemplo, Ignacio Aguado, miembro del Comité Ejecutivo de Ciudadanos o la cuenta del partido en Andalucía.

Figura 6. Panel de sugerencias de Twitter



Fuente: Elaboración propia

Aun así, ha habido usuarios que no aparecían y se ha tenido que indagar fuera de Twitter para detectar aquellos actores importantes dentro del ámbito político del partido para posteriormente realizar la búsqueda en la plataforma, dónde se ha decidido si el usuario era relevante o no en función de una serie de criterios que se explicarán más tarde.

2. En paralelo, se ha recogido, a través de un script en Python y haciendo uso de la librería tweepy, a los representantes de cada partido para las elecciones generales de abril de 2019. Esta información es pública en Twitter y su autor es el usuario @dalvarez37 que, a través de listas en Twitter, publica una serie de datos significativos para política española.

Este proceso se ha repetido por cada uno de los cinco grandes partidos en España: PSOE, PP, Ciudadanos, Unidas Podemos y VOX.

Para el resto de los partidos (ERC, JxCAT, EHBildu, CCanaria, Compromís, EAJPNV y EnMarea) se ha realizado tan solo la fase de 'Mejora de la red con nuevos usuarios relevantes' debido, a priori, a la menos representatividad en el esquema político español.

7.3. Criterios de elaboración

A lo largo del proceso de elaboración de la red, se han ido tomando determinadas decisiones en función de una serie de criterios:

- El proceso seguido para la elaboración de las subredes de ERC, JxCAT, EHBildu, CCanaria, Compromís, EAJPNV y EnMarea difiere del resto de partidos debido, en primera instancia, a su menor representatividad y por lo tanto relevancia en la red política española.

Es importante aclarar que más de uno de esos partidos interviene activamente y con poder en diferentes partes del territorio español. La decisión se tomó debido a que estos partidos eran relevantes, sobre todo, en solo una comunidad autónoma. Aunque en el trabajo se han podido tocar aspectos referentes a elecciones municipales, no era el objetivo principal de este y, por lo tanto, no se les ha considerado de gran relevancia en vista a las elecciones generales.

Otro de los motivos ha sido la diferencia de volumen en cuanto a representantes políticos de estos partidos en comparación con los cinco más importantes.

Por último, el número de actores externos de cierta relevancia de estas subredes (medios de comunicación, periodistas, etc.), dista mucho de los otros cinco principales partidos.

- El principio seguido durante el proceso de descarte de usuarios ha sido:
 - o Se descartarán todos aquellos usuarios cuyo número de seguidores esté por debajo de los 100, una cifra lo suficientemente baja como para que cualquier usuario de la plataforma en activo pueda alcanzarlo. De esta forma, nos desharemos de aquellas cuentas ‘fake’ de otras oficiales que tanto su éxito como su número de seguidores acaban siendo bajos. Aun así, no es una norma estricta y está abierta a interpretación. Podría suceder que en el momento de la detección de usuarios relevantes, algún futuro integrante clave de la red se cree una cuenta en Twitter y aún no haya logrado esa cifra. En ese caso, también pasaría a formar parte de la red.
 - o Se descartarán aquellos usuarios que representasen a cuentas de partidos políticos en ciudades no relevantes. Se ha definido un límite de 200.000 habitantes para determinar la relevancia de una ciudad. Por ejemplo, las cuentas locales como la de Unidas Podemos en Alcoi no serán tenidas en cuenta, pero sí la de Madrid o Barcelona. Sin embargo, vuelven a existir excepciones como, por ejemplo, Pamplona, justo por debajo del límite pero considerada una de las ciudades más importantes de España, si sería parte de la red.

- La tendencia seguida durante el proceso de selección de usuarios relevantes para la fase de ‘Ampliación y mejora de la red con nuevos usuarios relevantes’ ha sido navegar a través de usuarios en Twitter comenzando por la cuenta oficial del partido político en España o la de su líder.

El modo de transitar de un usuario a otro ha sido a través de las sugerencias que establece Twitter para cada usuario. Suelen ser usuarios que mantienen una estrecha relación con el usuario en cuestión en Twitter. Se ha completado este proceso con los criterios establecidos en el punto anterior.

7.4. Tipología de usuarios

Para la elaboración de la red política española no sólo se han tenido en cuenta usuarios en Twitter que representasen a políticos de los diferentes partidos en España. También se han considerado otro tipo de usuarios. Son los siguientes:

- Medios de comunicación: independientemente del tipo (prensa, radio, televisión, etc.). Por ejemplo, Antena3 (@antena3com, medio de comunicación televisivo) o El País (@el_pais, medio de comunicación de prensa). Este tipo de usuarios se han tenido en cuenta debido a la estrecha relación que existe entre muchos de ellos con el ámbito político y dado que interesa, para nuestro caso de estudio, su conducta a lo largo de una serie de eventos de carácter político.
- Líderes y representantes: tanto el líder de partido como representantes de este. Por ejemplo, Pablo Iglesias (@Pablo_Iglesias_, Secretario General de Unidas Podemos) o Juan Pedro Yllanes (@jpyllanes, diputado al Congreso por las Islas Baleares). Como es lógico, la red política de cualquier naturaleza estará formada por políticos.
- Cuentas institucionales: cuentas que representan al partido político en diferentes territorios de España. Por ejemplo, Podemos (@ahorapodemos, cuenta oficial del partido) o Podemos en la comunidad de Madrid (@PodemosCMadrid). Al igual que con los políticos, parece clara la inclusión de las cuentas institucionales de partidos en la red.
- Organizaciones – Activistas: organizaciones de diverso carácter relacionadas al partido en cuestión. Por ejemplo, Open Arms, organización sin ánimo de lucro dedicada a la protección de aquellas personas que huyen hacia Europa fruto de conflictos políticos en su país de origen (@openarms_fund, cuenta oficial de la organización Open Arms) o cuentas de movimientos reivindicativos como @M15Mbizkaia, cuenta oficial del movimiento 15-M en Bizkaia. La gran mayoría de estas organizaciones acaban adquiriendo una imagen política debido a la naturaleza de su actividad. Por ejemplo, en el caso de Open Arms, su función se ve muy relacionada con la polémica existente en España y, en Europa en general, sobre inmigración. Por ello, se ha considerado relevante introducir a ciertas organizaciones o activistas, siempre que cumplan los criterios del apartado 7.3.
- Periodistas – Comunicadores: Periodistas o aquellos que directamente realicen la función de comunicador a través de un medio de comunicación. Por ejemplo, Jordi Évole (@jordievole, cuenta oficial del periodista en La Sexta, Jordi Évole)

o Antonio Maestre (@AntonioMaestre, cuenta oficial del periodista colaborador en La sexta o Telemadrid, Antonio Maestre). La implicación de ciertos periodistas en política es alta, en definitiva, forman parte de los medios de los que ya hemos hablado.

- Programas de TV – Radio: programas de los medios de comunicación recogidos. Por ejemplo, ‘Salvados’ (@salvadostv, cuenta oficial del programa de televisión ‘Salvados’ en La Sexta) o ‘La Vida Moderna’ (@vidamoderna, cuenta oficial del programa de radio ‘La Vida Moderna’ en La Ser). Al igual que con los periodistas y comunicadores, son una parte de más que componen lo que conocemos como medio de comunicación. Por ese motivo, se ha decidido integrar este tipo de usuarios en la red. Es importante recalcar que solo se han tenido en cuenta aquellos que posean cierta relación con temas políticos, independientemente de la seriedad o formalidad de estos. Por ejemplo, el programa ‘La vida moderna’ es un programa de humor, pero puede llegar a tener más actividad e influencia en Twitter que un programa de contenido estrictamente político.
- ‘Influencers’: usuarios que participan activamente en Twitter y poseen un gran poder de influencia en sus seguidores. Por ejemplo, @protestona1. Existen cuentas que se dedican exclusivamente a su actividad en la plataforma y actúan, sobre todo, en temas políticos. Muchos de ellos también se autodefinen de un partido o de otro. En este caso, @protestona1 es afín a Unidas Podemos y lo deja claro en cada tweet. Este tipo de usuarios llegan incluso a tener más influencia que ciertas cuentas institucionales o representantes políticos, ya que rápido suelen hacerse eco de noticias negativas sobre la oposición y no tardan en compartirlo a través de la plataforma. Además suelen ser usuarios muy activos durante todo el día. Son razones suficientes como para formar parte de la red.
- Artistas: artistas cercanos al entorno del partido político en cuestión. Por ejemplo, la actriz Anabel Alonso (@AnabelAlonso_of, cuenta oficial de la actriz Anabel Alonso). Es menos común que el resto de usuarios, pero existen casos en los que actores se posicionan claramente hacia una ideología o partido político. En este caso, la actriz Anabel Alonso actúa activamente cercana al círculo de Unidas Podemos. Pero no es el único motivo por el que se puede posicionar a un usuario en un cuarto partido político. También pueden formar parte de la red por el simple hecho de participar con un alto grado de actividad sobre temas políticos.

Debemos incidir en que la distribución de estos usuarios, salvo líderes, representantes y cuentas institucionales, en su mayoría no se posicionan estrictamente sobre un partido político. Se trata de una asignación a priori, en función de la cercanía de sus relaciones con el partido político que trate.

Lo importante hasta este punto es definir la red completa para su posterior análisis, dónde se podrán observar como existen cambios sobre el posicionamiento de algunos actores.

8. Recogida de datos en Twitter

Para la consecución de los objetivos planteados en el trabajo, se ha considerado necesaria la recolección de información en Twitter en ciertos eventos de carácter político para la detección de comportamientos relevantes dentro de la red política española previamente elaborada.

Los eventos elegidos para la detección de información en Twitter han sido tres.

El primer evento es el debate desarrollado en Radio Televisión Española el día 22 de abril de 2019 a las 22 horas, con una previa de una hora. En él participaron 4 fuerzas políticas: PSOE, PP, Ciudadanos y Unidas Podemos. A VOX no se le permitió participar por no tener representación parlamentaria en las anteriores elecciones generales. El debate duró un total de 100 minutos en el que los cuatro partidos compartieron con el espectador sus propuestas divididas en cuatro grandes bloques:

1. Política económica, fiscal y empleo
2. Política social, estado del bienestar, pensiones e igualdad
3. Política territorial
4. Regeneración democrática y pactos postelectorales

El segundo evento es el debate emitido en Atresmedia el día 23 de abril de 2019 a las 22 horas con una previa de una hora. En él, al igual que en el de Radio Televisión Española, participaron los cuatro grandes partidos políticos, no permitiendo participar a VOX por el mismo motivo. El debate duró dos horas y los líderes de cada partido tuvieron que ofrecer su postura en base a tres grandes bloques, a la vez que debían gestionar un tiempo que el programa otorgaba a cada candidato de manera equitativa. Los tres temas que propuso Atresmedia fueron:

1. Programa electoral
2. Modelo territorial
3. Pactos

El tercer evento es el propio día de las elecciones, es decir, el 28 de abril de 2019.

Durante esta fase se ha utilizado Python como lenguaje para la extracción y MongoDB como herramienta de almacenamiento de esta. Sin embargo, este proceso se puede dividir en dos subprocesos según la tipología en la extracción de datos.

8.1. Streaming API

El primer subproceso es Streaming API de Twitter. Lo que se consigue es abrir una conexión entre nuestra aplicación y la API de Twitter de modo que, en función de los parámetros de búsqueda, devuelva información en tiempo real. En otras palabras, devuelve tweets que cumplan dicho criterio.

Para iniciar la conexión se ha hecho uso de las credenciales privadas como desarrolladores que proporciona Twitter que se observan en la Figura 7.

Figura 7. Credenciales de desarrollador en Twitter

Keys and tokens

Keys, secret keys and access tokens management

Consumer API keys

6EjOaE50qj5B71AxCjbUciMpg (API key)

27QfLDp2gdt8pZmG72P4bfGtYEPOWshsqiraomw0siwfbO6uTa (API secret key)

Regenerate

Access token & access token secret

1101126661918081025-3RCXtObC0S9TVMGjq5yp4T4ITgomlv (Access token)

I5E8L4M6huWDK9NBAX5gPaK8UJzJiwpxHD2bYy7KjtlqA (Access token secret)

Read and write (Access level)

Revoke

Regenerate

Fuente: Elaboración propia

Estas credenciales han sido almacenadas en un script para hacer uso de ellas siempre que fuese necesario sin necesidad de declararlas en cada código. De esta forma, solo con importar el script y sus variables, se puede tener acceso a la API de Twitter.

En la Figura 8 podemos ver el script.

Figura 8. Script de credenciales

```
# Variables that contains the user credentials to access Twitter API
consumer_key = "6EjOaE50qj5B[REDACTED]"
consumer_secret= "27QfLDp2gdt8pZmG72P4bfGtYEPOWshsqiraomw0siwfbO6uTa[REDACTED]"
access_token= "1101126661918081025-3RCXtObC0S9TVMGjq5yp4T4ITgomlv[REDACTED]"
access_token_secret= "I5E8L4M6huWDK9NBAX5gPaK8UJzJiwpxHD2bYy7KjtlqA[REDACTED]"
```

Fuente: Elaboración propia

El criterio de búsqueda establecido es que recogiese todo tweet que involucrase a cualquier usuario de nuestra red. Esto incluye tweets, respuestas, menciones, citas y retweets a alguno de sus tweets, aunque la acción la realice un usuario que no hayamos definido en el criterio de búsqueda, pero sí el del tweet originario. Para ello, ha sido necesario el id de cada uno de los usuarios de nuestra red.

Para esto último se necesitó crear otro script en Python, conocer los nombres de usuario de cada actor de la red y la ayuda de tweepy. Todo ello, en este punto, ya se tendría.

Los nombres de usuario ya han sido recogidos durante la elaboración de la red en hojas de cálculo. El siguiente paso es trasladar todos esos nombres a una lista. Para ello se ha utilizado un archivo de texto donde poder dar formato a través de la sustitución de caracteres, ya que cada nombre debía aparecer doblemente entrecomillado en la lista.

Una vez se tuvo la lista preparada, se debió crear el script para hacer llamadas a la API de Twitter y recibir, en función del nombre de usuario, su id (número identificativo en Twitter), correspondiente.

Es importante incidir en que solo se permiten hacer llamadas para un máximo de doscientos usuarios. Por lo tanto, la lista se ha dividido en varias listas más pequeñas, en total nueve.

La lógica seguida en el código se observa en la Figura 9. Es decir, por cada usuario definido en la lista, obtener el nombre e id del usuario a través de la función `get_user` que facilita la API de twitter.

Figura 9. Script que recibe el id en Twitter en función del nombre de usuario

```
while True:
    try:
        for usuario in usuarios9:
            user=api.get_user(usuario)
            print(usuario,user.id)

    except tweepy.TweepError as e:
        print(e)
        print("El usuario ", usuario, " no ha sido encontrado")
        continue
```

Fuente: Elaboración propia

El siguiente paso fue ir almacenando los números identificativos en un archivo para crear una lista de id's.

Una vez se tuvo esta lista, se pudo poner en funcionamiento el script de escucha. La lógica seguida se puede entender en las Figuras 10, 11 y 12.

En la Figura 10 se puede ver como se importa el script de credenciales, las librerías `tweepy`, para el proceso de escucha, `json` para dar formato a la información, `pymongo`

para acceder a la base de datos MongoDB y urllib3.exception y http.client para tratar diferentes excepciones en caso de error.

El método get_auth permite autenticar al usuario con sus correspondientes claves de acceso.

Figura 10. Parte 1 del script Streaming API

```
from secret import consumer_key, consumer_secret, access_token, access_token_secret
import tweepy
import json
from pymongo import MongoClient
from urllib3.exceptions import ProtocolError
from http.client import IncompleteRead

MONGO_HOST= 'mongodb://localhost/twitterdb' # assuming you have mongoDB installed locally
                                             # and a database called 'twitterdb'

def get_auth():
    auth = tweepy.OAuthHandler(consumer_key, consumer_secret)
    auth.set_access_token(access_token, access_token_secret)
    return auth
```

Fuente: Elaboración propia

Ha sido necesario crear una clase en la que se definió el comportamiento del script. En la Figura 11 se observa el método on_connect que establece la conexión con la API de Twitter. Por otro lado, on_error avisa en caso de error. Por último, on_data crea una base de datos en MongoDB, parsea la información a formato json y acaba insertando cada tweet que viene en una colección dentro de la base de datos. En ambos casos, si la base de datos o la colección no existen, las crea automáticamente.

Figura 11. Parte 2 del script Streaming API

```
class MyStreamListener(tweepy.StreamListener):

    def on_connect(self):
        # Called initially to connect to the Streaming API
        print("You are now connected to the streaming API.")

    def on_error(self, status_code):
        # On error - if an error occurs, display the error / status code
        print('An Error has occured: ' + repr(status_code))
        return False

    def on_data(self, data):
        #This is the meat of the script...it connects to your mongoDB and stores the tweet
        try:
            client = MongoClient(MONGO_HOST)

            # Use twitterdb database. If it doesn't exist, it will be created.
            db = client.twitterdb

            # Decode the JSON from Twitter
            datajson = json.loads(data)

            #grab the 'created_at' data from the Tweet to use for display
            created_at = datajson['created_at']

            #print out a message to the screen that we have collected a tweet
            print("Tweet collected at " + str(created_at))

            #insert the data into the mongoDB into a collection called twitter_search
            #if twitter_search doesn't exist, it will be created.
            db.elecciones.insert_one(datajson)

        except Exception as e:
            print (e)
```

Fuente: Elaboración propia

En la Figura 12 se aprecia cómo se construye el objeto API y el 'Listener' de la clase creada para referenciarlos con `tweepy` e iniciar el proceso de escucha.

Con `myStream.filter` se define el criterio de selección. En este caso, se desean todos los tweets en los que interaccionen de una forma u otra los usuarios de la red definida previamente y reflejada en la lista 'miRed'.

Figura 12. Parte 3 del script Streaming API

```
# Get an API item using tweepy
auth = get_auth() # Retrieve an auth object using the function 'get_auth' above
api = tweepy.API(auth) # Build an API object.

# Connect to the stream
myStreamListener = MyStreamListener()
myStream = tweepy.Stream(auth=api.auth, listener=myStreamListener)

miRed=["87818409", "145992645", "82084630", "22954354", "2561091", "79939248", "762402774260875265"]

print(">> Listening to tweets about _____:")
while True:
    try:
        myStream.filter(follow=miRed)
    except (ProtocolError, AttributeError, IncompleteRead) as e:
        continue
```

Fuente: Elaboración propia

Es importante aclarar que el script ha sido ejecutado desde una hora antes de los debates de RTVE y Atresmedia hasta una hora y media después de estos.

Esto significa que el día 22 de abril de 2019, se ejecutó el script desde las 22:02 hasta la 1:00 del 23 de abril de 2019, este rango de tiempo hace referencia al período de tiempo que abarcó el debate en RTVE desde el comienzo del debate hasta una hora después de la finalización de este.

Para el debate de Atresmedia se siguió la misma lógica, pero abarcando una hora más, lo que significaba escuchar la previa del debate. El script funcionó desde las 21:05 hasta la 1:00 del día siguiente.

Por último, el día de las elecciones, el 28 de abril de 2019, el script permaneció a la escucha desde las 14:07 hasta la 1:38 del día siguiente.

8.2. API REST

El API Rest de Twitter también permite obtener información de Twitter, pero no en tiempo real, sino información pasada. En concreto, se puede recoger información con hasta siete días de antigüedad.

Igualmente posee otra diferencia respecto al streaming API, en vez de establecer una conexión continua, en este caso se entregan los datos en lotes a través de solicitudes repetidas de nuestra aplicación.

Para iniciar dichas llamadas también se han utilizado las credenciales. En común con el streaming, también se puede fijar el criterio de búsqueda. En este caso se han especificado una serie de hashtags, en concreto, lo más seguidos en los días de las elecciones y los debates. Para que la información adquiriese cierto valor también se ha fijado un período de búsqueda. En este caso, como ejemplo, se ha buscado sobre #28A entre los días 28 y 29 de abril.

Es importante aclarar que el proceso inicial de autenticación con credenciales y conexión a base de datos con pymongo es idéntico al anterior script.

En la Figura 13 se puede ver como se define un criterio de búsqueda. Lo primero que se especifica son los hashtags a seguir con el parámetro 'q', en este caso en relación a los debates de Radio Televisión Española y la cadena Atresmedia.

En segundo lugar, se define el rango de tiempo en el que se quiere realizar la búsqueda con los parámetros 'since' y 'until' para referirse a desde y hasta cuando respectivamente. Como es natural, coincide con los días en los que se celebraron dichos debates.

A continuación, por cada tweet que cumplía dicho criterio, se ha almacenado en una colección, en este caso llamada 'filtroPorHashtagsMartes'.

Por último, si se producía un error, el sistema descansaba tres minutos y continuaba por donde ocurrió el error, en lugar de terminar o reiniciar el proceso.

Figura 13. Script API REST

```
while True:
    try:
        for tweet in tweepy.Cursor(api.search,
                                   q=("##DebateRTVE OR #DebateAtresmedia OR #ElDebateDecisivo OR #ElDebateEnRTVE"),
                                   count=100,since="2019-04-23",
                                   until="2019-04-24",
                                   tweet_mode= 'extended').items():
            db.filtroPorHashtagsMartes.insert_one(tweet._json)

    except tweepy.TweepError as e:
        if 'Failed to send request:' in e.reason:
            print("Time out error caught")
            time.sleep(180)
            continue
```

Fuente: Elaboración propia

Para el debate de RTVE, los hashtags utilizados fueron: [#DebateRTVE](#) y [#ElDebateEnRTVE](#).

Para el debate de Atresmedia, los hashtags utilizados fueron: [#DebateAtresmedia](#) y [#ElDebateDecisivo](#).

Para el día de las elecciones, los hashtags utilizados fueron: [#28A](#), [#HoyUnidasPodemos](#), [#MásPPMásEspaña](#), [#VamosCiudadanos](#), [#28APSOE](#), [#PorEspaña](#), [#OlaNaranja](#), [#EleccionesGenerales28Abril](#), [#HazQuePase](#), [#EleccionesGenerales](#) y [#EleccionesGenerales28A](#).

8.3. Búsquedas complementarias

Además de los dos tipos de obtención de datos en Twitter, se ha necesitado un tercer proceso complementario para la obtención de los ‘follows’ de cada usuario de la red previamente elaborada. Es decir, a quien sigue cada usuario de la red.

Esto ha servido como fase de construcción del grafo para el análisis estático en el que se ha representado la relación de los usuarios en función de quién sigue a quién.

Para ello se ha creado un script como se ve en la Figura 14, en el que se almacenaron en la colección ‘follows’ los archivos con el número identificativo y la lista de ‘follows’ por cada usuario de la red.

Figura 14. Script de recogida y almacenamiento de ‘follows’

```
def save_to_mongo(custom_object):

    client = MongoClient()
    db = client.follows_user

    db.follows.insert_one(custom_object).inserted_id

def get_follows_user(usuario):
    amigosDeMiRed = []
    friends_ids = []
    num_seguidos_mi_red=0
    num_seguidos=0
    try:
        for friend_id in tweepy.Cursor(api.friends_ids, id=usuario).items():
            friends_ids.append(friend_id)
            num_seguidos+=1
        print("Los seguidos por parte de: ",usuario," son:\n ",friends_ids)
        print(num_seguidos)
        print("-----")
        custom_object = {
            "screen_name": usuario,
            "follows": friends_ids,
        }

        save_to_mongo(custom_object)
    except tweepy.TweepError as tweep_error:
        print("Error with code : " + str(tweep_error.response.text))

for usuario in miRed:
    get_follows_user(usuario)
```

Fuente: Elaboración propia

Posteriormente, con otro script se ha completado el proceso para disponer de una colección con el nombre de usuario, número identificativo y lista de ‘follows’ por cada actor que compone la red, como se puede ver en la Figura 15.

Con este código se recorre la colección anterior y se realiza una petición por cada usuario en la que se obtiene el nombre de usuario, creando así una colección con los tres atributos mencionados.

La obtención del nombre de usuario (`screen_name` en el código) es vital para la comprensión de la red en su visualización. Sería casi imposible detectar determinados usuarios en un grafo viendo tan solo su número identificativo.

Figura 15. Script para la creación de un objeto personalizado con ‘follows’ de cada usuario

```
def save_to_mongo(custom_object):  
  
    client = MongoClient()  
    db = client.follows_user  
    db.followsDeMiRedConNombre.insert_one(custom_object).inserted_id  
  
def filtrado():  
    for tweet in tweets:  
        user=api.get_user(tweet['screen_name'])  
        custom_object={  
            "id": tweet['screen_name'],  
            "screen_name": user.screen_name,  
            "followsMyNet": tweet['followsMyNet'],  
        }  
        save_to_mongo(custom_object)  
  
filtrado()
```

Fuente: Elaboración propia

9. Almacenamiento

Para el almacenamiento de la información extraída de Twitter se ha utilizado la herramienta MongoDB. Es un sistema de base de datos NoSQL orientado a documentos de código abierto. En lugar de guardar los datos en tablas, tal y como se hace en las bases de datos relacionales, MongoDB guarda estructuras de datos BSON (una especificación similar a JSON) con un esquema dinámico, haciendo que la integración de los datos sea más fácil y rápida. Por esto último, se ha optado por MongoDB y no por otra base de datos.

La gestión de los datos se puede realizar desde consola o con la herramienta complementaria MongoDB Compass, que es más intuitiva.

Como se ha comentado en el apartado anterior, para tener acceso a MongoDB se debe hacer uso de la librería pymongo.

Si se observa la Figura 16, MongoClient permite hacer referencia al host utilizado, en este caso 'mongodb://localhost/twitterdb'.

Figura 16. Fragmento de código para el acceso a la base de datos MongoDB

```
from pymongo import MongoClient

MONGO_HOST= 'mongodb://localhost/twitterdb' # assuming you have mongoDB installed locally
                                                # and a database called 'twitterdb'
```

Fuente: Elaboración propia

Para añadir información (tweets en este caso) a la correspondiente colección, se ha hecho uso del fragmento de código de la Figura 17.

En el código, db hace referencia a la base de datos donde se encuentra la colección donde se quiere insertar determinada información.

Figura 17. Referencia a la base de datos en MongoDB

```
db.filtroPorHashtagsMartes.insert_one(tweet._json)
```

Fuente: Elaboración propia

Para ello, se ha definido previamente con el código de la Figura 18.

Figura 18. Declaración de la base de datos en MongoDB

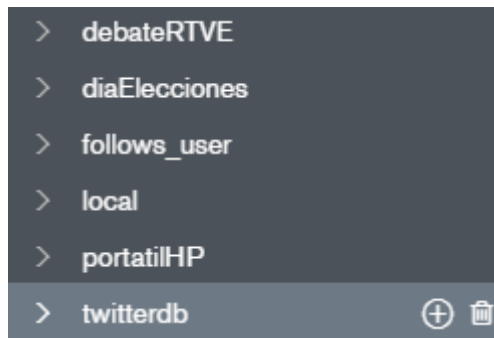
```
db = client.twitterdb
```

Fuente: Elaboración propia

Por lo tanto, en este caso, se ha insertado un nuevo tweet en la colección ‘filtroPorHashtagsMartes’ en la base de datos ‘twitterdb2’.

En MongoDB Compass, se puede observar la base de datos ‘twitterdb’ de las 10 existentes (Figura 19).

Figura 19. Aspecto en MongoDB Compass de las bases de datos



Fuente: Elaboración propia

Dentro de la base de datos se puede encontrar varias colecciones, como se muestra en la Figura 20.

Figura 20. Aspecto en MongoDB Compass de las colecciones de una base de datos

Collection Name ^	Documents	Avg. Document Size	Total Document Size	Num. Indexes	Total Index Size
debateA3Media	85,099	5.9 KB	494.2 MB	1	900.0 KB
filtroPorHashtags1	642,994	5.6 KB	3.5 GB	1	6.0 MB
filtroPorHashtagsLunes2	950,571	5.7 KB	5.1 GB	1	9.0 MB
filtroPorHashtagsMartes	584,249	5.7 KB	3.2 GB	1	5.4 MB

Fuente: Elaboración propia

Como se puede apreciar, aparece la colección ‘filtroPorHashtagsMartes’. Dicha colección consta de 584.249 tweets y ocupa un total de 3.2 GB.

Cada tweet sigue la estructura que aparece en la Figura 21.

Figura 21. Estructura por campos de cada tweet

```
_id: ObjectId("5cbf662079ce6ba212497f74")
created_at: "Tue Apr 23 19:23:07 +0000 2019"
id: 1120770091732414464
id_str: "1120770091732414464"
text: "RT @eljueves: Es posible morir de vergüenza ajena oyendo el minuto de ..."
source: "<a href='\"http://twitter.com/download/iphone\"' rel='\"nofollow\"'>Twitter fo..."
truncated: false
in_reply_to_status_id: null
in_reply_to_status_id_str: null
in_reply_to_user_id: null
in_reply_to_user_id_str: null
in_reply_to_screen_name: null
> user: Object
  geo: null
  coordinates: null
  place: null
  contributors: null
> retweeted_status: Object
  is_quote_status: false
  quote_count: 0
  reply_count: 0
  retweet_count: 0
  favorite_count: 0
> entities: Object
  favorited: false
  retweeted: false
  filter_level: "low"
  lang: "es"
  timestamp_ms: "1556047387427"
```

Fuente: Elaboración propia

Es sustancial la diferencia en la estructura de la información recogida entre Streaming API y API Rest. En la siguiente imagen se puede ver la estructura de un tweet recogido con Streaming API.

La estructura de la información recogida a través de API Rest no incluye algunos campos anteriores. Estos son ‘text’, ‘quote_count’, ‘reply_count’, y ‘timestamp’. Pero incluye otros nuevos como ‘full_text’, ‘display_text_range’, ‘metadata’, y ‘possibly_sensitive’.

MongoDB Compass también proporciona un menú de filtrado de tweets según el criterio establecido. Es una herramienta útil para comparar a primera vista determinadas colecciones, por ejemplo, se puede utilizar para observar si dos colecciones tienen el mismo objetivo de recogida o no. Esto es debido a que se han realizado varias pruebas de recogida para aumentar la fiabilidad de lo que se estaba extrayendo. Por lo tanto, un modo sencillo de saberlo era ordenar las colecciones por el campo ‘timestamp’ desde el principio y ver si la información coincidía o no.

Por último, la recogida de datos se ha hecho con dos equipos en paralelo debido a que, en ocasiones, el script podría errar por la ingente cantidad de datos que llegaba en tiempo real.

De este modo, reducíamos la pérdida de datos en caso de que esto ocurriese. Por esto, ha sido necesaria la exportación e importación de algunas colecciones en concreto a otros equipos.

Hay dos formas de realizar esta operación:

1. Desde consola:

- Para exportar, con la siguiente estructura:

```
mongoexport --db test --collection traffic --out traffic.json
```

Donde 'db' indica la base de datos, 'collection' la colección y 'out' el nombre del archivo resultante de la exportación.

- Para importar, con la siguiente estructura:

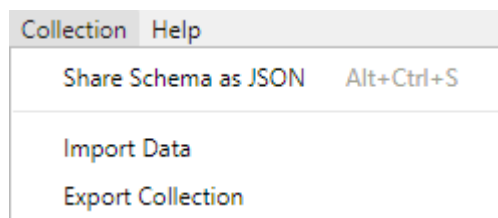
```
mongoimport --db users --collection contacts --file contacts.json
```

Donde 'db' indica la base de datos, 'collection' la colección 'file' el nombre del archivo que se quiere importar.

2. Desde MongoDB Compass:

Tanto para la exportación como para la importación, desde el menú desplegable izquierdo superior, se ofrecen ambas opciones (Figura 22).

Figura 22. Menú desplegable izquierdo superior en MongoDB Compass



Fuente: Elaboración propia

En un primer momento el proceso se realizó desde consola, pero el proceso era bastante más lento, así que se acabó utilizando MongoDB Compass como herramienta para la exportación e importación de colecciones.

10. Procesado de datos

Una vez terminada la fase de recolección y almacenamiento de datos en Twitter, se necesitó procesar esa información para disponer tan solo de la relevante, que sirvió en la posterior creación y análisis de grafos.

El siguiente paso fue filtrar los ‘follows’ de los usuarios de la red, de modo que solo se obtuvieran aquellos ‘follows’ que perteneciesen a la propia red. Es decir, se consiguen los usuarios de la red que siguen a otros usuarios de esta. Por ejemplo, si se dispone de la lista de ‘follows’ de Pablo Iglesias, solo se necesitan aquellos ‘follows’ que formen parte de la red política española previamente elaborada.

Para ello, se ha creado un script como se puede ver en la Figura 23.

Con este código, recorreremos la colección creada en pasos anteriores y si el número identificativo está en nuestra lista de usuarios ‘miRed’, lo añade a una nueva colección. De este modo solo se obtienen aquellos usuarios que pertenecen a nuestra red y a su vez son ‘follow’ de algún integrante de esta.

Figura 23. Script de filtrado de ‘follows’ en la red

```
def save_to_mongo(custom_object):

    client = MongoClient()
    db = client.follows_user

    db.followsDeMiRed.insert_one(custom_object).inserted_id

def filtrado():
    for tweet in tweets:
        for follow in tweet['follows']:
            if follow in miRed:
                friends_idsMiRed.append(follow)
    custom_object={
        "id":tweet['id'],
        "screen_name": tweet['screen_name'],
        "followsMyNet": friends_idsMiRed,
    }
    save_to_mongo(custom_object)

filtrado()
```

Fuente: Elaboración propia

El siguiente paso fue juntar todas aquellas colecciones que hacían referencia a un mismo evento. Para ello, el proceso fue muy sencillo, tan solo había que recorrer cada colección e ir

añadiendo cada tweet a una colección común. En la Figura 24 se observa el proceso seguido. En este caso, por cada tweet de la colección ‘porHashtagEleccionesDomingo5Mitad2’ de la base de datos ‘twitterdb’ se inserta en la colección ‘diaElecciones’ de la base de datos ‘diaElecciones’. La idea es seguir este proceso con todas las colecciones almacenadas referentes al mismo evento.

Figura 24. Script para creación de colecciones

```
db = client.twitterdb
db2=client.diaElecciones
tweets = db.porHashtagEleccionesDomingo5Mitad2.find()

def filtrado():
    for tweet in tweets:
        db2.diaElecciones.insert(tweet, check_keys=False)
```

Fuente: Elaboración propia

Hasta este punto se han conseguido 1.133.022 tweets publicados el día del debate de Radio Televisión Española, 1,447,360 tweets para el debate de Atresmedia y 1,136,150 del día de las elecciones.

Antes de entrar a la visualización de los datos, era necesario eliminar los tweets duplicados. Para ello, hay una forma bastante sencilla que proporciona MongoDB. En este caso se ha realizado por consola, como se puede apreciar en la Figura 25.

Donde el parámetro ‘laClave’ es el atributo por el cual se decide si existe un elemento en la colección idéntico o no. El segundo parámetro tenido en cuenta es ‘coleccion’ donde definiremos la colección donde se quiere realizar la eliminación de duplicados.

Figura 25. Borrado de duplicados en MongoDB por el campo ‘id’

```
> db.diaElecciones.aggregate([
...   { "$group": {
...     "_id": { "id": "$id" },
...     "dups": { "$push": "$_id" },
...     "count": { "$sum": 1 }
...   }},
...   { "$match": { "count": { "$gt": 1 } }}
... ], { allowDiskUse: true }).forEach(function(doc) {
...   doc.dups.shift();
...   db.diaElecciones.remove({ "_id": {"$in": doc.dups } });
... });
```

Fuente: Elaboración propia

El número de tweets almacenados finalmente fueron 731.419 en relación con el debate de Atresmedia, 669.280 en relación con el debate de Radio Televisión Española y 876,407 del día de las elecciones generales, el 28 de abril.

Pero, no todos los tweets eran útiles, ya que la mayoría no eran tweets, retweets, respuestas o citas producidas por ningún usuario de la red. Por este motivo se debía realizar un filtrado de los tweets, de modo que se obtuviesen solo aquellos en los que interactuasen activamente los usuarios de nuestra red.

Esta operación no se expone en este apartado ya que involucra la parte de visualización y creación de grafos. Ambos procesos se han producido en paralelo.

11. Construcción de grafos

Una vez se poseía la información relevante, el siguiente paso fue la construcción de los grafos. Esta fase del proyecto se realizó con Python y la librería NetworkX y se limitó a la creación de los grafos, sin entrar en visualizar. Se consideró así porque, aunque se puede visualizar el grafo resultante con NetworkX, no es especialmente eficiente para gran cantidad de elementos en la red. Para la visualización de los grafos se utilizará otra herramienta.

Esta parte se puede dividir en dos.

La primera parte consiste en la creación del grafo objetivo para la realización del análisis estático.

Para este proceso se ha seguido la lógica que se muestra en el script de la Figura 26. Lo primero de todo, fue necesario definir un grafo dirigido con la cláusula `DiGraph()` para observar en la etapa de visualización en qué dirección se produce la relación. El siguiente paso fue la creación de un nodo por cada usuario de la red.

Figura 26. Creación del grafo y nodos iniciales

```
DG=nx.DiGraph()
DG.add_nodes_from(['87818409','145992645','82084630','22954354']
)
```

Fuente: Elaboración propia

Una vez estén creados todos los nodos, se dio paso a la creación de relaciones entre ellos.

Para ello se ha empleado el fragmento de código de la Figura 27.

Por cada usuario guardado en la colección creada en pasos previos, si cada follow de su lista de follows ('followMyNet' en el código) coincide con algún usuario de la red, se crea una relación, representada en forma de arista.

Figura 27. Creación de relaciones en función de los 'follows' en la red

```
def filtrado():
    for tweet in tweets:
        for follow in tweet['followsMyNet']:
            if str(follow) in miRed:
                DG.add_edge(tweet['screen_name'],str(follow))

filtrado()
```

Fuente: Elaboración propia

De tal modo, se ha obtenido un grafo dirigido, es decir, mostrando la dirección de la relación establecida. También se puede declarar al grafo como binario, lo que significa que la relación entre dos nodos no tiene valor, solo existe o no. Por otro lado, es inconexo, ya que existen diez nodos que no siguen ni son seguidos por ningún otro usuario de la red. Por último, la red resultante es unidireccional debido a que solo mantienen un tipo de relación.

Por otro lado, se necesita la creación de uno o varios grafos para el análisis dinámico. Estos grafos se construyeron en función de las interacciones en los tres eventos mencionados en el apartado 8.

En este caso, el grafo varía en cierta medida respecto al anterior, ya que las relaciones entre nodos pasan a tener un valor, llamado peso. Por lo tanto, este grafo es valorado.

En la Figura 28, la clase 'nodes_creation' crea un nodo en el grafo por cada usuario de la red elaborada. El nodo creado tiene como atributos el número identificativo, nombre de usuario, número de seguidores, número de seguidos, número de tweets publicados y fecha de creación de la cuenta. Es importante remarcar que el primer atributo del nodo será la etiqueta cuando se represente, aunque más tarde pueda cambiarse.

Figura 28. Creación de nodos inicial para el grafo dinámico

```
def nodes_creation():
    i=coleccionDeUsers.count()
    while True and i>0:
        try:
            for tweet in coleccionDeUsers:
                user=api.get_user(tweet['screen_name'])
                print("Vamos por el usuario ",tweet['screen_name'])
                DG.add_node(tweet['screen_name'], id_str = tweet['id'], seguidores=user.followers_count,
                            seguidos=user.friends_count, tweets= user.statuses_count,
                            creacion=str(user.created_at))
            i-=1
        except Exception as e:
            print(e,"----- EL USUARIO ", tweet['screen_name'], "NO HA SIDO ENCONTRADO")
            print(i)
            i-=1
            continue
```

Fuente: Elaboración propia

Una vez creados los nodos, es hora de crear las relaciones en función de las interacciones.

El código se puede dividir en tres funcionalidades, como se puede observar en la Figura 29

La primera se basa en la detección de citas en Twitter. Para ello se observa si quien realiza la cita y el citado aparecen en nuestra lista de usuarios de la red.

En caso afirmativo, crea una relación representada con una arista que une ambos nodos. En caso negativo continua con el siguiente tweet de la colección. En caso de que ya exista esa

relación, se le suma 0.5 al valor del peso. De esta manera creamos un grafo capaz de representar la fortaleza de las relaciones entre nodos.

La segunda funcionalidad consta del mismo proceso, pero con la detección de retweets. Si el usuario que realiza el retweet y el autor del tweet retweeteado aparecen en nuestra lista de usuarios en la red, se crea la relación.

Por último, la tercera consiste en la misma operación, pero en este caso con la detección de respuestas.

Figura 29. Creación de relaciones para el grafo dinámico

```
def all_interactions():
    try:
        print("-----")
        print("Comienza la creación de QUOTES!")
        for tweet in coleccion:
            if ('quoted_status' in tweet) and (tweet['quoted_status']['user']['id_str'] in miRed)
            and (tweet['user']['id_str'] in miRed):
                if DG.has_edge(tweet['user']['id_str'], tweet['quoted_status']['user']['id_str']):
                    DG[tweet['user']['id_str']][tweet['quoted_status']['user']['id_str']]['weight'] += 0.5
                else:
                    DG.add_edge(tweet['user']['id_str'], tweet['quoted_status']['user']['id_str'], weight = 0.5)
                    print("arista creada entre ", tweet['user']['screen_name'],
                        " y ", tweet['quoted_status']['user']['screen_name'])
        print("-----")
        print("Comienza la creación de RETWEETS!")
        for tweet in coleccion3:
            if ('retweeted_status' in tweet) and (tweet['retweeted_status']['user']['id_str'] in miRed)
            and (tweet['user']['id_str'] in miRed):
                if DG.has_edge(tweet['user']['id_str'], tweet['retweeted_status']['user']['id_str']):
                    DG[tweet['user']['id_str']][tweet['retweeted_status']['user']['id_str']]['weight'] += 0.5
                else:
                    DG.add_edge(tweet['user']['id_str'], tweet['retweeted_status']['user']['id_str'], weight = 0.5)
                    print("arista creada entre ", tweet['user']['screen_name'],
                        " y ", tweet['retweeted_status']['user']['screen_name'])
        print("-----")
        print("Comienza la creación de RESPUESTAS!")
        for tweet in coleccion4:
            if (tweet['in_reply_to_screen_name']!=null) and (tweet['in_reply_to_user_id_str'] in miRed)
            and (tweet['user']['id_str'] in miRed):
                if DG.has_edge(tweet['user']['id_str'], tweet['in_reply_to_user_id_str']):
                    DG[tweet['user']['id_str']][tweet['in_reply_to_user_id_str']]['weight'] += 0.5
                else:
                    DG.add_edge(tweet['user']['id_str'], tweet['in_reply_to_user_id_str'], weight = 0.5)
                    print("arista creada entre ", tweet['user']['screen_name'],
                        " y ", tweet['in_reply_to_screen_name'])
```

Fuente: Elaboración propia

12. Representación gráfica

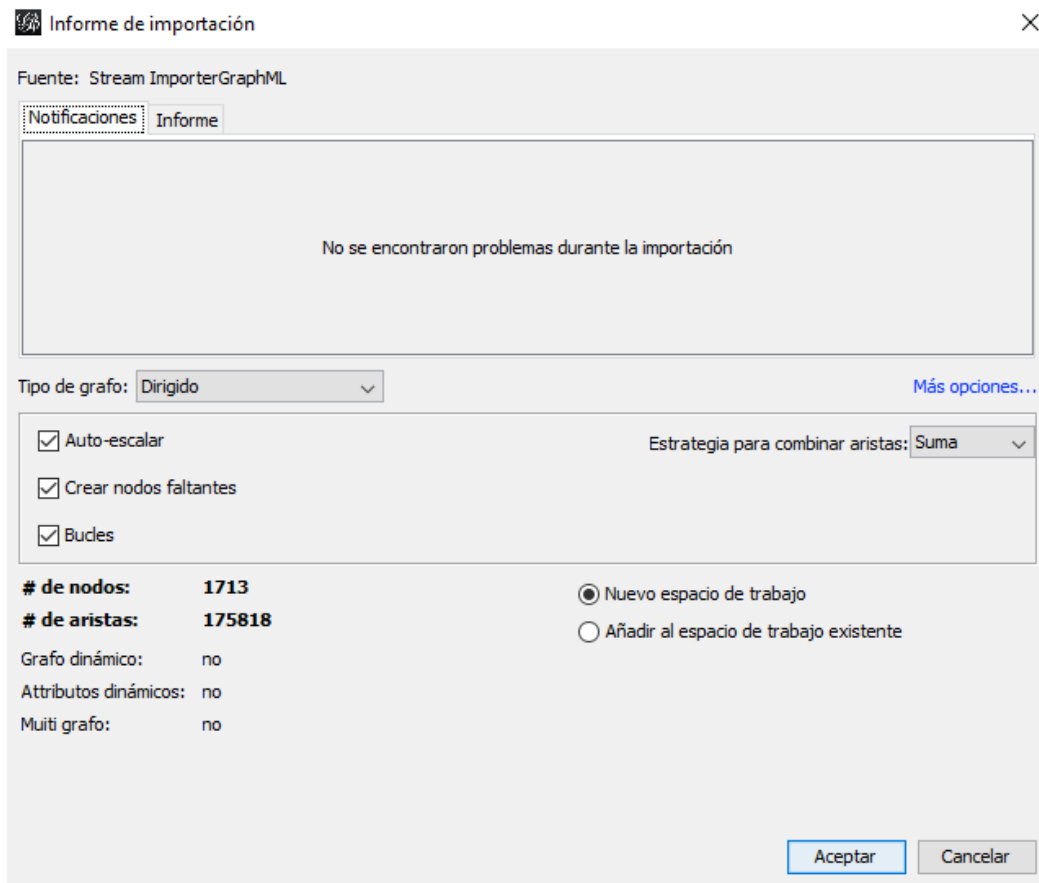
Una vez creados los grafos, se dio comienzo a la fase de visualización. En esta parte del trabajo se ha utilizado Gephi como herramienta.

Gephi ha permitido comprender mejor la estructura de grafos y la localización de patrones en estos a través de la manipulación de la red. Además, ha permitido la elaboración de conclusiones a partir de singularidades dentro de la red.

Durante esta parte se han utilizado una serie de herramientas en las diferentes etapas del proceso de visualización del grafo.

La primera etapa consiste en la importación del set de datos a Gephi.

Figura 30. Panel de importación de Gephi



Fuente: Elaboración propia

En este caso, se ha importado el grafo estático creado a partir de la relación de usuarios en función de si se siguen o no.

Como se aprecia en la imagen, se puede definir si el grafo es dirigido o no. En este caso sí lo es, de hecho, recordemos que se ha construido como grafo dirigido con NetworkX.

En la Figura 31 se aprecia el estado inicial del grafo tras su importación.

Muestra un estado poco valorable, por lo que se debe comenzar a utilizar las diferentes herramientas que proporciona Gephi.

Figura 31. Estado inicial del grafo

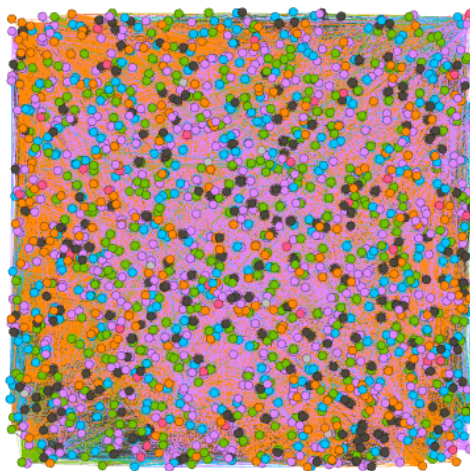


Fuente: Elaboración propia

El siguiente paso, seguido a lo largo del trabajo, ha sido hallar aquellas comunidades que detecta. Para ello se puede definir una resolución determinada, en este caso se ha dejado la que viene por defecto, uno.

Una vez detectadas las comunidades, se ha dado color a cada nodo en función de la comunidad a la que pertenece, como se observa en la Figura 32.

Figura 32. Cambio de color de los nodos en función de la modularidad



Fuente: Elaboración propia

Gephi facilita una serie algoritmos de distribución.

Durante el trabajo, los más usados han sido:

- Force Atlas 2: aunque es idóneo para redes pequeñas, va a ser utilizado para la primera toma de contacto con la red estática, la más grande durante el trabajo. Con este algoritmo se ha observado como los elementos se organizan en función de su comunidad. En el centro del grafo aparecen los elementos más conectados, mientras los menos conectados están en zonas periféricas del grafo.
- Fruchterman-Reingold: simula la acción de elementos metálicos y fuerzas eléctricas que se ejercen sobre ellos. Acaba generando un grafo circular donde se logra detectar con suficiencia cada comunidad. Ha sido utilizado en subredes más pequeñas ya que no es eficiente para redes con gran cantidad de elementos.
- OpenOrd: es idóneo para la detección de comunidades. Ha sido utilizado para contrastar los resultados ofrecidos por Force Atlas 2, pero no como resultado final.

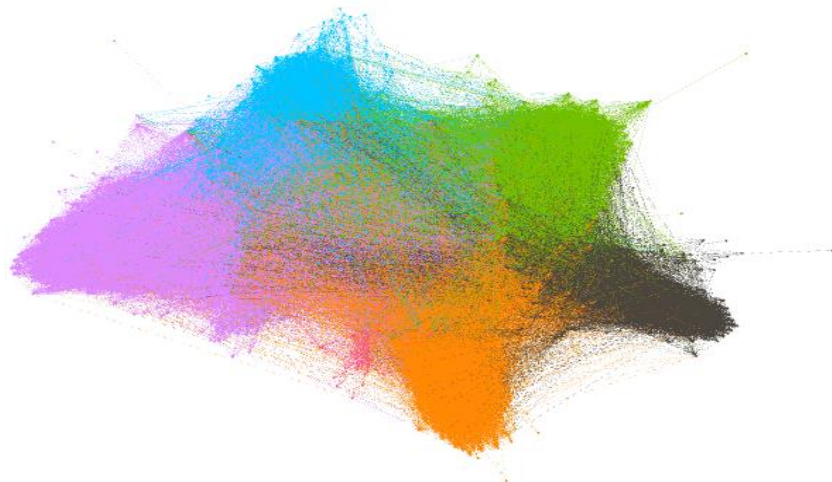
También ofrece algoritmos de ajuste como, por ejemplo:

- Noverlap: expande los nodos. Durante el trabajo se ha utilizado para observar con más detalle cada comunidad tras aplicar Force Atlas 2 o OpenOrd.
- Ajuste de etiquetas: cambia progresivamente la localización de los nodos para evitar el solapamiento de etiquetas.

Una vez explicado cada algoritmo utilizado durante el proceso se puede entender mejor qué se ha realizado paso a paso.

Si se continúa con el grafo anterior, el primer paso ha sido aplicar el algoritmo Force Atlas 2. Ahora se puede detectar con más exactitud cada comunidad en la Figura 33. A priori, parece que hay cinco: azul, morado, naranja, negro y verde.

Figura 33. Grafo tras aplicar el algoritmo Force Atlas

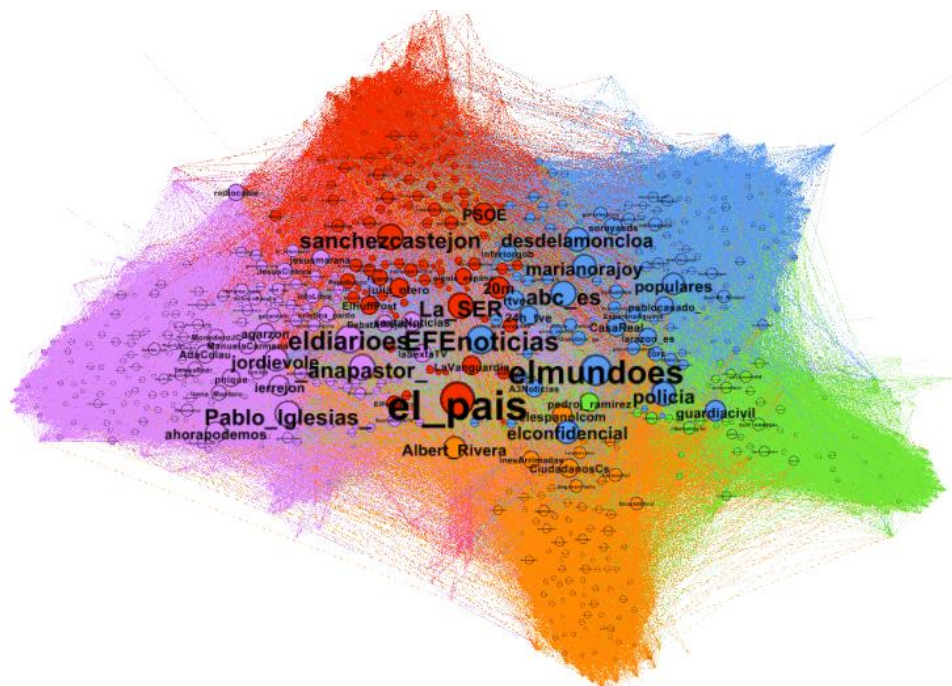


Fuente: Elaboración propia

El siguiente paso fue la visualización de los usuarios. Para ello, ha habido que permitir la visualización de las etiquetas y dar un tamaño en función de la relevancia del usuario en la red. En este caso se ha utilizado como criterio el número de seguidores en la red. El número de seguidores está representado por el grado de entrada de cada nodo. Este último dato se ha generado en el apartado de Estadísticas ejecutando el grado medio.

En la Figura 34 se ve como las etiquetas están solapadas. Si se ejecuta el ajuste de etiquetas y se aplica un color representativo a cada partido político del grafo, se obtiene una imagen más clarificadora como la de la Figura 35.

Figura 35. Grafo tras el ajuste de etiquetas



Fuente: Elaboración propia

Para el caso de comunidades más pequeñas como la subred de partidos independentistas, se ha seguido la misma metodología, pero supliendo Force Atlas 2 por Fruchterman-Reingold. El resultado tras aplicar este algoritmo es algo similar a la figura 36.

13. Cálculo de la polaridad en la red

Por último, en paralelo con la construcción del grafo, se ha creado una nueva colección en la que se han almacenado las interacciones entre usuarios en los diferentes eventos. Estos documentos incluyen origen y destino de la interacción, peso, texto y polaridad del tweet. Esto último fue de ayuda para el apartado del caso de estudio, donde ha servido como punto de apoyo para sustentar lo planteado en él.

Para ello, se ha seguido la lógica reflejada en la Figura 37.

Como ya se ha visto, dependiendo del método de recolección de tweets, aparecen unos campos u otros. De esta forma observamos si debemos tomar el campo 'full_text' o 'text'.

Gracias a la librería TextBlob, se puede obtener el grado de polaridad de un mensaje, en este caso, el campo de texto de cada tweet. TextBlob es una librería de Python que permite el procesamiento de textos. Gracias a la API que proporciona se es capaz de llevar a cabo la clasificación de textos a través del análisis de sentimiento.

Esta herramienta devuelve el sentimiento ('sentiment' en Python) de un texto. Este campo está formado a su vez por otros dos; estos campos son la polaridad y la subjetividad.

Figura 37. Cálculo de la polaridad

```
if ('full_text' in tweet):
    tweetLimpio=clean_tweet(tweet['full_text'])
    analisis=TextBlob(tweetLimpio)
    analisis = analisis.sentiment
    popularidad = analisis.polarity
```

Fuente: Elaboración propia

La polaridad es la emoción expresada y la subjetividad es cuando el texto es explicativo o muestra alguna opinión o creencia y debe ser analizado en contexto. En este caso, se ha optado por recoger tan solo aquellos tweets por debajo del 0,1 de subjetividad durante los tres eventos explicados en el apartado 15, tratando de ser lo más objetivo posible para no distorsionar el resultado final. Una vez se tiene la polaridad del tweet, se crea una variable que agrupa los campos ya citados como en la Figura 38.

Figura 38. Creación de una variable personalizada de polaridad

```
custom_object={
    "origen":tweet['user']['screen_name'],
    "destino":tweet['quoted_status']['user']['screen_name'],
    "peso":0.5,
    "polaridad":popularidad,
    "texto":tweetLimpio
}
```

Fuente: Elaboración propia

Una vez creado, solo es necesario guardarlo en una nueva colección:

db2.relacionesRTVE.insert_one(custom_object)

A partir de esta colección se ha creado un grafo (apartado 11) en función de las relaciones establecidas. Este grafo ha sido visualizado en Gephi y posteriormente, a través del laboratorio de datos, se ha exportado la información a una hoja de cálculo, dónde se ha llevado a cabo un procesado de la información.

Este procesado trata de reunir toda la atención en aquellas relaciones que conlleven un grado de polaridad. De estas relaciones se obtuvieron los usuarios que participan con cierto grado de polaridad. Por otro lado, se ha dividido la polaridad en dos tipos según la posición del usuario en la interacción.

Si el usuario actúa como nodo de origen, el grado de polaridad se ha definido comportamiento, ya que es el grado de sentimiento con el que un determinado usuario interactúa con el resto.

Por otro lado, si el usuario es el nodo de destino, el grado de polaridad hace referencia a la popularidad de este, puesto que es cómo el resto de usuarios interactúan con él, reflejando así, como ven al referido.

Por último, se ha dividido este conjunto de datos en los tres eventos con los que se ha trabajado a lo largo del proyecto, para poder estudiar la evolución de los resultados obtenidos.

14. Análisis estático de la red

14.1. Estructura

A continuación, se presenta el resultado de la red tras su elaboración.

La red política española resultante, lograda a través del proceso de elaboración, consta de un total de 1730 usuarios. Del total de usuarios para cada partido, de mayor a menor número de usuarios, los diferentes partidos políticos cuentan con:

- Unidas Podemos con 435 usuarios (25,14% de la red)
- PSOE con 330 usuarios (19,07% de la red)
- VOX con 310 usuarios (17,91% de la red)
- PP con 304 usuarios (17,57% de la red)
- Ciudadanos con 240 usuarios (13,87% de la red)
- ERC con 25 usuarios (1,44% de la red)
- JxCat con 25 usuarios (1,44% de la red)
- EHBildu con 13 usuarios (0,75% de la red)
- Coalición Canaria con 13 usuarios (0,75% de la red)
- EAJPNV con 12 usuarios (0,69% de la red)
- Compromís con 12 usuarios (0,69% de la red)
- En Marea con 11 usuarios (0,63% de la red)

Del total de usuarios para cada tipo de cuenta, de mayor a menor número de usuarios, cada tipo cuenta con:

- 751 líderes o representantes de partido (43,41% de la red)
- 425 cuentas institucionales (24,56% de la red)
- 215 organizaciones o activistas (12,42% de la red)
- 126 medios de comunicación (7,28% de la red)
- 89 periodistas o comunicadores (5,14% de la red)
- 58 influencers (3,35% de la red)
- 37 programas de radio o televisión (2,13% de la red)
- 9 artistas (0,52% de la red)

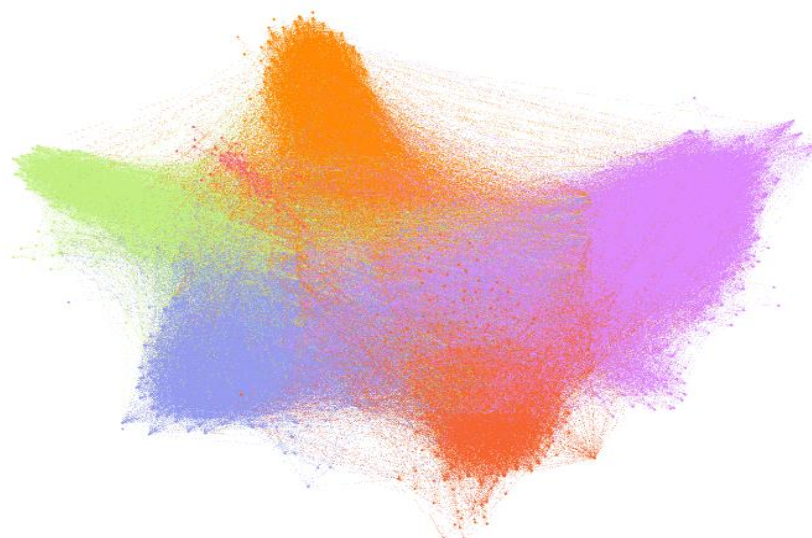
Teniendo en cuenta solo aquellos usuarios que podemos posicionar de forma precisa (cuentas institucionales y líderes o representantes de partido), se observan distinciones entre los partidos políticos:

- Cuentas institucionales:
 - 109 Unidas Podemos (25,64%)
 - 89 Ciudadanos (20,94%)
 - 88 PP (20,70%)
 - 79 VOX (18,58%)
 - 41 PSOE (9,64%)
 - 19 Cuentas comunes de gobierno (p. e., cuentas de ministerios) (4,47%)
- Líderes o representantes de partido:
 - 169 PP (25,64%)
 - 165 PSOE (25,03%)
 - 153 Unidas Podemos (23,21%)
 - 118 Ciudadanos (17,90%)
 - 54 VOX (8,19%)

Para completar el análisis estático de la red, se ha creado un grafo en función de que usuarios siguen a otros usuarios de esta. Así, se ha podido ver como se establecen las relaciones entre los distintos actores de la red.

Para su análisis, se ha visualizado el grafo de diferentes formas en los próximos apartados, dando mayor notoriedad a usuarios según los criterios que se utilicen. En la Figura 39 se puede ver como se estructura la red.

Figura 39. Red política española



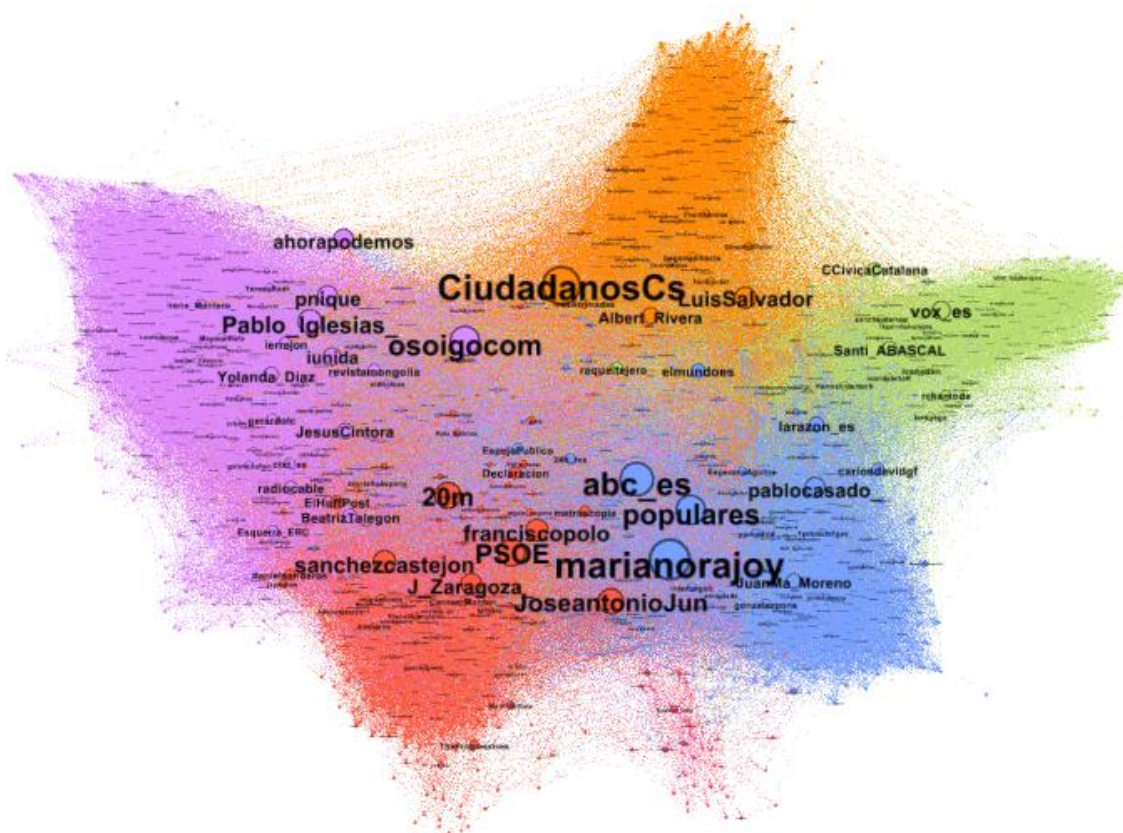
Fuente: Elaboración propia

14.2. Centralidad

Se puede medir la centralidad de la red de distintas formas.

En la Figura 40, los actores más destacados son aquellos que poseen mayor centralidad en función de criterio de centralidad de intermediación, es decir, el grado en que un nodo concreto forma parte de la distancia geodésica entre otros actores. Esto implica que para que dos actores mantengan relación, este actor debe ser un intermediario para que esto ocurra. Estos actores tienen cierto poder ya que controlan, de alguna manera, las interacciones entre pares de nodos.

Figura 40. Red política española en función de la centralidad de intermediación



Fuente: Elaboración propia

Los tres actores más destacados son @marianorajoy (PP), @CiudadanosCs (Ciudadanos) y @abc_es (posicionada en PP).

Si se entra en detalle, se pueden obtener los 3 usuarios más relevantes posicionado por cada partido:

En PSOE:

1. @PSOE: cuenta oficial del Partido Socialista Obrero Español en España.
2. @20m: cuenta oficial del periódico 20minutos.
3. @JoseAntonioJun: cuenta oficial del diputado por Granada del PSOE.

En Partido Popular:

1. @marianorajoy: cuenta oficial del expresidente del Partido Popular (2011 – 2018).
2. @abc_es: cuenta oficial del periódico ABC.
3. @populares: cuenta oficial del Partido Popular en España.

En Ciudadanos:

1. @CiudadanosCs: cuenta oficial de Ciudadanos en España.
2. @Luis_Salvador: cuenta oficial del portavoz de Ciudadanos en el ayuntamiento de Granada.
3. @Albert_Rivera: cuenta oficial del líder y máximo representante de Ciudadanos en España.

En Unidas Podemos:

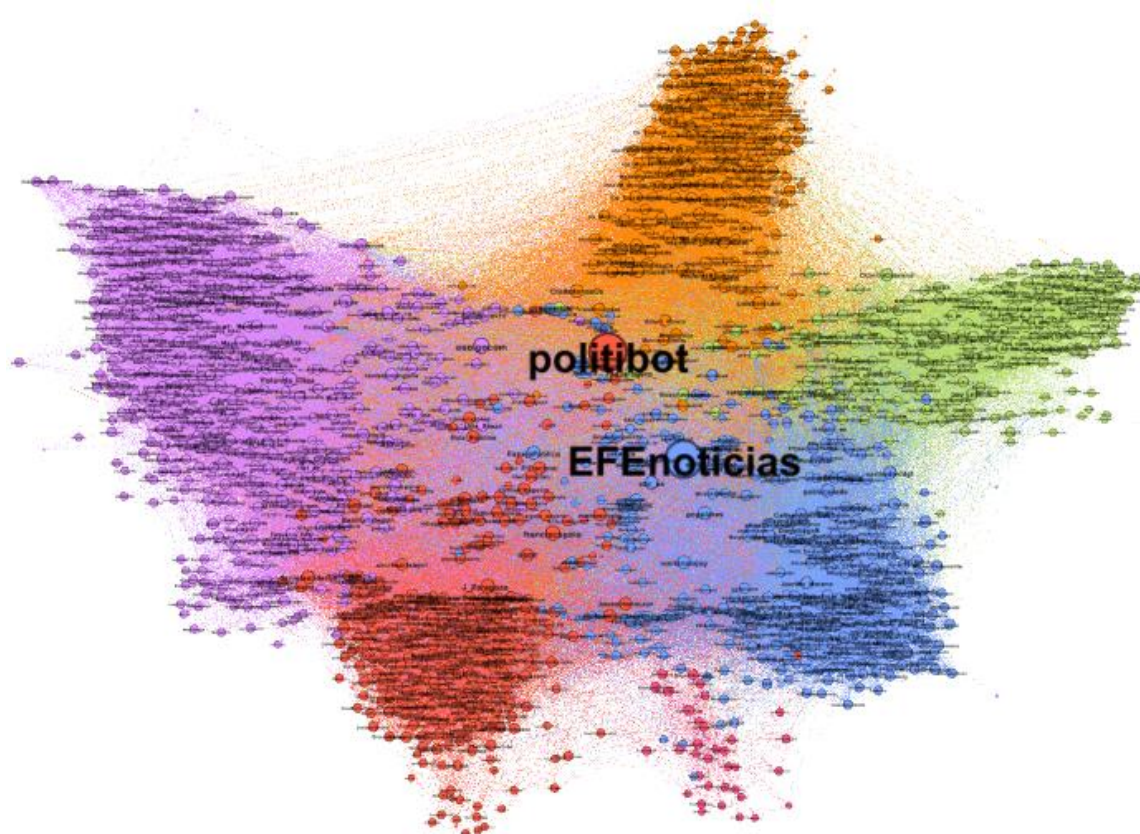
1. @osoigocom: cuenta oficial del proyecto de interacción entre ciudadanos y políticos a través de preguntas y respuestas.
2. @Pablo_Iglesias_: líder y máximo representante de Unidas Podemos en España.
3. @pniq: cuenta oficial del secretario de Organización de Unidas Podemos en España.

En VOX:

1. @vox_es: cuenta oficial de VOX en España.
2. @Santi_ABASCAL: líder y máximo representante de VOX en España.
3. @CCivicaCatalana: cuenta oficial de la organización Convivencia Cívica Catalana.

En la Figura 41, los actores más destacados son aquellos que poseen mayor centralidad en función del criterio de centralidad de cercanía, es decir, cómo de cerca está un nodo del resto. Mide la importancia estratégica de la red. Cuanto más central sea un nodo, más capacidad de interacción tendrá con el resto de los usuarios en la red.

Figura 41. Red política española en función de la centralidad de cercanía



Fuente: Elaboración propia

Los tres actores más destacados son @EFEnoticias (posicionada en PP), @politibot (posicionada en PSOE) y @osoigocom (posicionada en Unidas Podemos).

Si se entra en detalle, se pueden obtener los 3 usuarios más relevantes posicionados por cada partido:

En PSOE:

1. @politibot: cuenta dedicada a la detección de bulos en Twitter.
2. @franciscopolo: cuenta oficial del Secretario de Estado para el Avance Digital por el PSOE, Francisco Polo.
3. @J_Zaragoza_: cuenta oficial del diputado del PSC al Congreso, José Zaragoza.

En Partido Popular:

1. @EFEnoticias: cuenta oficial de la Agencia de noticias EFE.
2. @populares: cuenta oficial del Partido Popular en España.
3. @EspejoPúblico: cuenta oficial del programa informativo 'Espejo Público' en Antena3.

En Ciudadanos:

1. @CiudadanosCs: cuenta oficial de Ciudadanos en España.
2. @pepecanonucia: cuenta oficial del diputado en Cortes por Ciudadanos, Pepe Cano Fuster.
3. @LuisSalvador: cuenta oficial del portavoz de Ciudadanos en el ayuntamiento de Granada.

En Unidas Podemos:

1. @osoigocom: cuenta oficial del proyecto de interacción entre ciudadano y político a través de preguntas y respuestas.
2. @Yolanda_Diaz_: cuenta oficial de la candidata por Pontevedra en con En Común Podemos.
3. @pnique: cuenta oficial del secretario de Organización de Unidas Podemos en España.

En VOX:

1. @raqueltejero_: cuenta oficial de la periodista en OKDiario
2. @CCivicaCatalana: cuenta oficial de la organización Convivencia Cívica Catalana.
3. @eslamanadadeFJL: cuenta oficial del programa 'Es La Mañana de Federico', en esRadio.

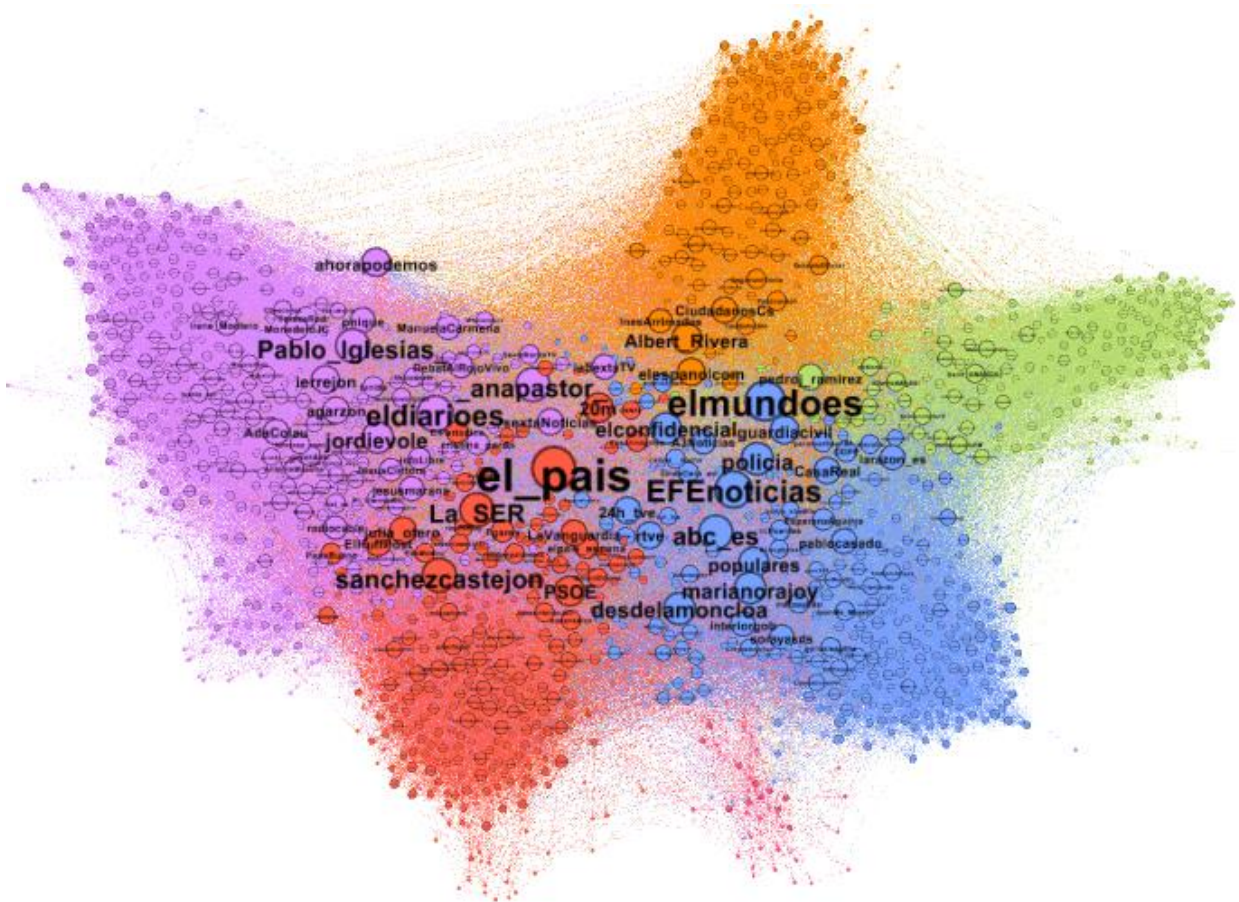
14.3. Prestigio

Se puede medir el prestigio de la red de diferentes formas.

La red presenta el siguiente aspecto:

En la Figura 42 los actores más destacados son los que más seguidores tienen dentro de nuestra red.

Figura 42. Red política española en función del número de seguidores



Fuente: Elaboración propia

Los tres actores más destacados son @el_pais (posicionada en PSOE), @elmundoes (posicionada en PP) y @EFEnoticias (posicionada en PP).

Si se entra en detalle, se pueden obtener los 3 usuarios más relevantes posicionados por cada partido:

En PSOE:

1. @el_pais: cuenta oficial del periódico 'El País'.
2. @sanchezcastejon: líder y máximo representante del PSOE en España.
3. @La_SER: cuenta oficial de la cadena de radio 'La SER'.

En Partido Popular:

1. @elmundoes: cuenta oficial del periódico 'El Mundo'.
2. @EFEnoticias: cuenta oficial de la Agencia de noticias EFE.
3. @abc_es: cuenta oficial del periódico 'ABC'.

En Ciudadanos:

1. @Albert_Rivera: cuenta oficial del líder y máximo representante de Ciudadanos en España.
2. @elespanolcom: cuenta oficial del periódico 'El Español'.
3. @CiudadanosCs: cuenta oficial de Ciudadanos en España.

En Unidas Podemos:

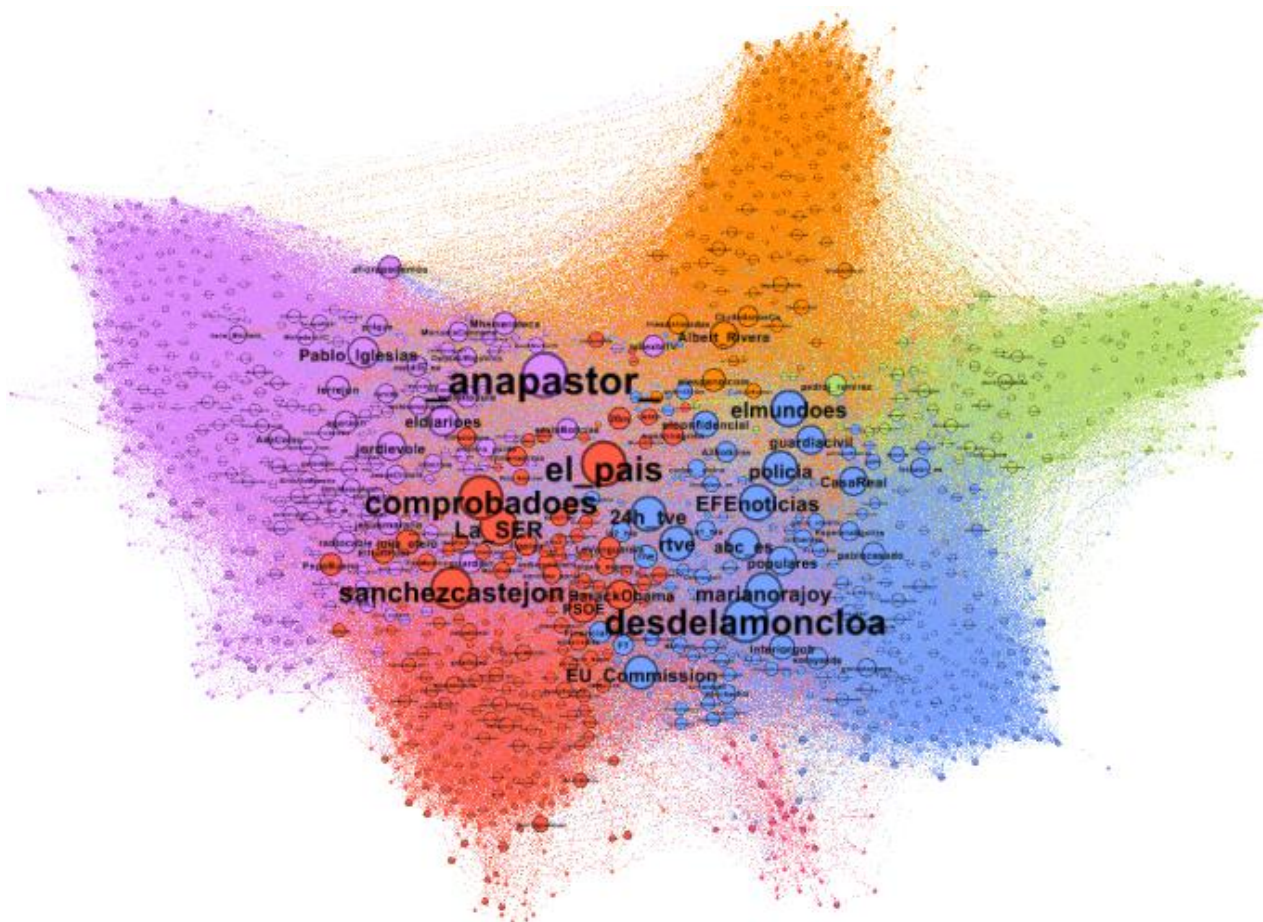
1. @eldiarioes: cuenta oficial del periódico digital 'El Diario'.
2. @_anapastor_: cuenta oficial de la periodista Ana Pastor.
3. @Pablo_Iglesias_: líder y máximo representante de Unidas Podemos en España.

En VOX:

1. @pedroj_ramirez: cuenta oficial del director del periódico 'El Español'.
2. @libertaddigital: cuenta oficial del periódico digital 'Libertad Digital'.
3. @Santi_ABASCAL: líder y máximo representante de VOX en España.

En la figura 43 los actores más destacados son los que más seguidores influyentes tienen dentro de nuestra red (Con el algoritmo Pagerank). El algoritmo Pagerank calcula la suma de los pesos de los enlaces entrantes tomando como valor el Pagerank del actor origen. Se utiliza la misma lógica que para las páginas en Internet, es decir, el valor del algoritmo tiene en cuenta la importancia de los nodos que se comunican con el nodo en cuestión. Así mismo, un usuario, a pesar de tener menos número de enlaces con otros usuarios, puede tener un alto grado de prestigio. De esta forma, existen entradas de enlaces más importantes que otras.

Figura 43. Red política española en función del Pagerank



Fuente: Elaboración propia

Por lo tanto, en este caso, la importancia resulta de la suma de las puntuaciones de Pagerank de los nodos que señalan al nodo en cuestión.

Los tres actores más destacados son @_anapastor_ (posicionada en Unidas Podemos), @desdelamoncloa (posicionada en PP) y @el_pais (posicionada en PSOE).

Si se entra en detalle, se pueden obtener los 3 usuarios más relevantes posicionados por cada partido:

En PSOE:

1. @el_pais: cuenta oficial del periódico 'El País'.
2. @comprobadoes: cuenta oficial de medios nacionales, regionales y locales que se unen con el objetivo de verificar el discurso público y político.
3. @sanchezcastejon: líder y máximo representante del PSOE en España.

En Partido Popular:

1. @desdelamoncloa: cuenta oficial del Gobierno de España.
2. @elmundoes: cuenta oficial del periódico 'El Mundo'.
3. @marianorajoy: cuenta oficial del expresidente del Partido Popular (2011 – 2018).

En Ciudadanos:

1. @Albert_Rivera: cuenta oficial del líder y máximo representante de Ciudadanos en España.
2. @elespanolcom: cuenta oficial del periódico 'El Español'.
3. @CiudadanosCs: cuenta oficial de Ciudadanos en España.

En Unidas Podemos:

1. @_anapastor_: cuenta oficial de la periodista Ana Pastor.
2. @Pablo_Iglesias_: líder y máximo representante de Unidas Podemos en España.
3. @eldiarioes: cuenta oficial del periódico 'El Diario'.

En VOX:

1. @pedroj_ramirez: cuenta oficial del director del periódico 'El Español'.
2. @Santi_ABASCAL: líder y máximo representante de VOX en España.

3. @libertaddigital: cuenta oficial del periódico digital ‘Libertad Digital’.

Si se suma el número de veces que un partido político está entre los 3 usuarios más relevantes según el criterio establecido, se obtiene lo siguiente:

1. PP (6 veces)
2. PSOE (3 veces)
3. Unidas Podemos (2 veces)
4. Ciudadanos (1 vez)

Si se toma como referencia el número de ocasiones en las que un usuario aparece entre los más relevantes de su partido, se puede detectar aquellos usuarios importantes a la par que poseen cierto grado de fiabilidad y consistencia en la red.

En el Partido Socialista, el usuario que más veces aparece entre los tres más relevantes es Pedro Sánchez (@sanchezcastejon), empatado con el periódico ‘El País’ (@elpais), con 2 apariciones.

En el caso del Partido Popular sería más complicado definir al usuario más importante, ya que se produce un empate a cinco entre Mariano Rajoy (@marianorajoy), la cuenta del Partido Popular (@populares), los periódicos ‘ABC’ y ‘El Mundo’ (@abc_es y @elmundoes respectivamente) y @EFEnoticias, una agencia de noticias internacional, todos con 2 apariciones.

Por el contrario, en Ciudadanos, destaca la cuenta oficial del partido (@CiudadanosCs) en solitario, con cuatro apariciones, seguido de Albert Rivera, con tres.

También, en Unidas Podemos, se distingue la cuenta de Pablo Iglesias (@Pablo_Iglesias_) con 3 apariciones.

Por último, en VOX, despunta el líder del partido Santiago Abascal (@Santi_ABASCAL) con 3 apariciones.

14.4. Subredes

A continuación, se puede ver un desglose por cada partido político en función del número de seguidores dentro de su propia subred. Los actores más destacados, son los que más seguidores tienen dentro de la subred de su partido.

En la subred de Unidas Podemos (Figura 44) se encuentran aquellos representantes y cuentas institucionales del partido, tanto de Podemos como de Izquierda Unida, así como los medios de comunicación afines a la ideología del partido como, por ejemplo, @eldiarioes.

Es destacable que, a pesar de la fragmentación del partido y creación de otros nuevos, actores como la alcaldesa de Madrid, Manuela Carmena y, el candidato a la Comunidad de Madrid por Más Madrid, Íñigo Errejón, se mantengan en la red de Unidas Podemos. Por otro lado, es algo normal ya que, al haber pertenecido al partido durante tanto tiempo, es difícil romper aquellas relaciones (follows en este caso) aun dejando por completo el partido. De hecho, incluso el partido autonómico Más Madrid acaba perteneciendo a la red de Unidas Podemos, algo comprensible dado la ideología de este y al estar formado por antiguos integrantes de Unidas Podemos. A pesar de no formar parte activamente del partido, disponen de una posición privilegiada en la red, ya que son el sexto y la décimo tercera con más seguidores dentro de la red

Es importante acentuar la aparición en la red de todos los partidos independentistas vascos y catalanes. Estos partidos son Esquerra Republicana (ERC), Junts per Catalunya (JxCat), Partido Nacionalista Vasco (PNV) y Euskal Herria Bildu (EHBildu). En cierto modo, es entendible el posicionamiento de estas subredes, dado que Unidas Podemos es el partido más cercano ideológicamente a la mayoría de estos. A pesar de que algunos de estos partidos independentistas sean considerados de derechas, gran culpa del posicionamiento lo tiene el conflicto catalán y vasco. Se puede observar como tiene más peso la postura respecto a la independencia que la ideología global de cada partido. Esto cobra sentido si se repara en que Unidas Podemos es la única fuerza que se ha posicionado a favor del diálogo con los partidos independentistas sobre este asunto.

Además, se observa como Compromís y EnMarea aparecen en la red de Unidas Podemos, algo esperado ya que ambos partidos se consideran de izquierdas y progresistas. De hecho, Podemos formaba parte de la antigua EnMarea, coalición de Podemos, EU, Equo Galicia, Esquerda Unida y Anova.

Se pueden encontrar programas de distinto tipo como, por ejemplo, el programa de radio 'La Vida Moderna' donde participan colaboradores como Juan Ignacio Delgado Alemany (@IgnatiusFarray) y Héctor de Miguel Martín (@_Queque_), también pertenecientes a esta subred.

También existen artistas en la red de Unidas Podemos como la actriz Anabel Alonso (@AnabelAlonso_of) igualmente conocida por su actividad en Twitter sobre temas de

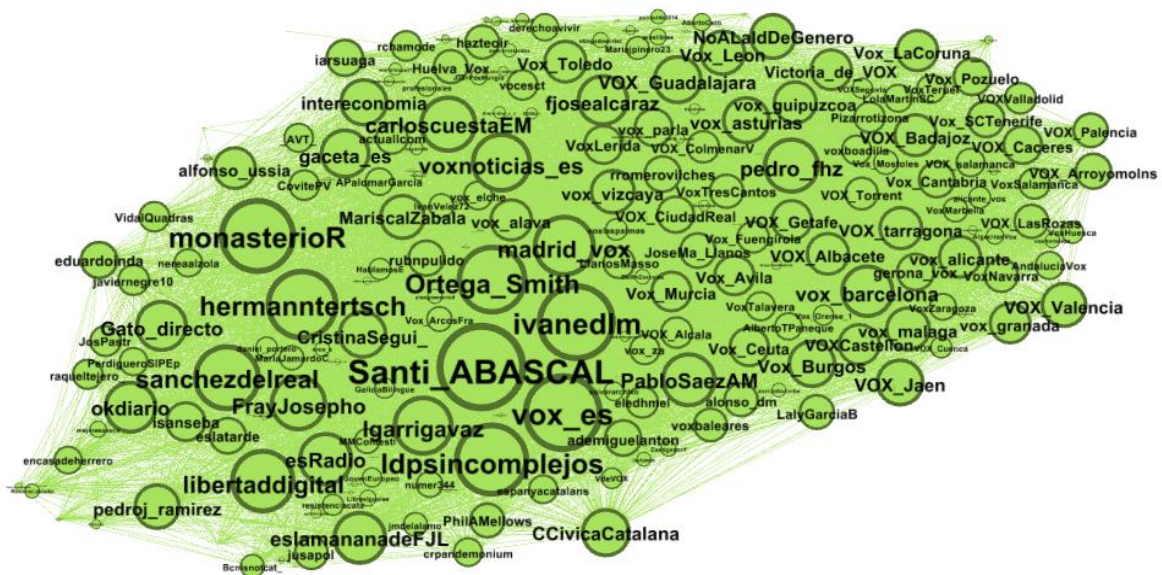
(@monasterioR), aunque no quedan muy distanciados nodos como los de Javier Ortega Smith (@Ortega_Smith) o Hermann Tertsch (@hermanntertsch).

Pero, sin duda, la singularidad de la red de VOX (Figura 48) es la fuerte comunidad de medios de comunicación y periodistas que respaldan a la red y hacen de hilo conductor con el resto de comunidades. Son muchos los ejemplos de medios afines a la ideología de VOX como, por ejemplo, Intereconomía (@intereconomia), ‘OkDiario’ (@Okdiario) o ‘Libertad Digital’ (@libertaddigital). También podemos encontrar comunicadores como Eduardo Inda (@eduardoinda), Cristina Seguí (@CristinaSegui_), Raquel Tejero (@raqueltejero_), Alfonso Ussia (@alfonso_ussia), Luis del Pino (@ldpsincomplejos) o Javier Negre (@javiernegre10).

Esta comunidad se ve reforzada con organizaciones con una tendencia ideológica cercana al entorno de VOX como, por ejemplo, HazteOir (@hazteoir), Convivencia Cívica Catalana (@CCivicaCatalana), Resistencia Catalana (@resistenciacata) y Derecho a Vivir (@derechoavivir).

Por último, la red de VOX cuenta con el apoyo de varios influencers con una cierta tendencia conservadora como @FrayJosepho, @JosPastr, @joveneuropeo, @mejoreszascas o @PhilAMellows. Junto con Unidas Podemos, es el único partido que posee en su red tal cantidad de integrantes en Twitter que no ejercen la política ni ocupan un cargo público, siendo en su mayoría usuarios anónimos.

Figura 48. Subred de VOX



Fuente: Elaboración propia

14.5. Alternativas en la elaboración de la red

Como ya se ha comentado, los partidos independentistas catalanes y vascos se posicionan en la red de Unidas Podemos, pero es entendible dada la situación política actual. Además, la lógica que sigue el algoritmo de modularidad en Gephi es agrupar nodos en función de la interacción, pero en ningún caso por su ideología política, aunque pueda coincidir. Por lo tanto, las comunidades creadas se rigen por algún tipo de criterio en común dada la interacción de estos. Es tarea nuestra interpretar los resultados. En este caso, existen varias posibilidades. La primera, como ya se ha mencionado, es que Unidas Podemos es el único partido que apuesta por el diálogo contra el conflicto catalán. Otra alternativa es la similitud que muestran estos partidos en Twitter; todos muestran una actitud más directa e informal en la plataforma, lejos de los partidos tradicionales como PP o PSOE.

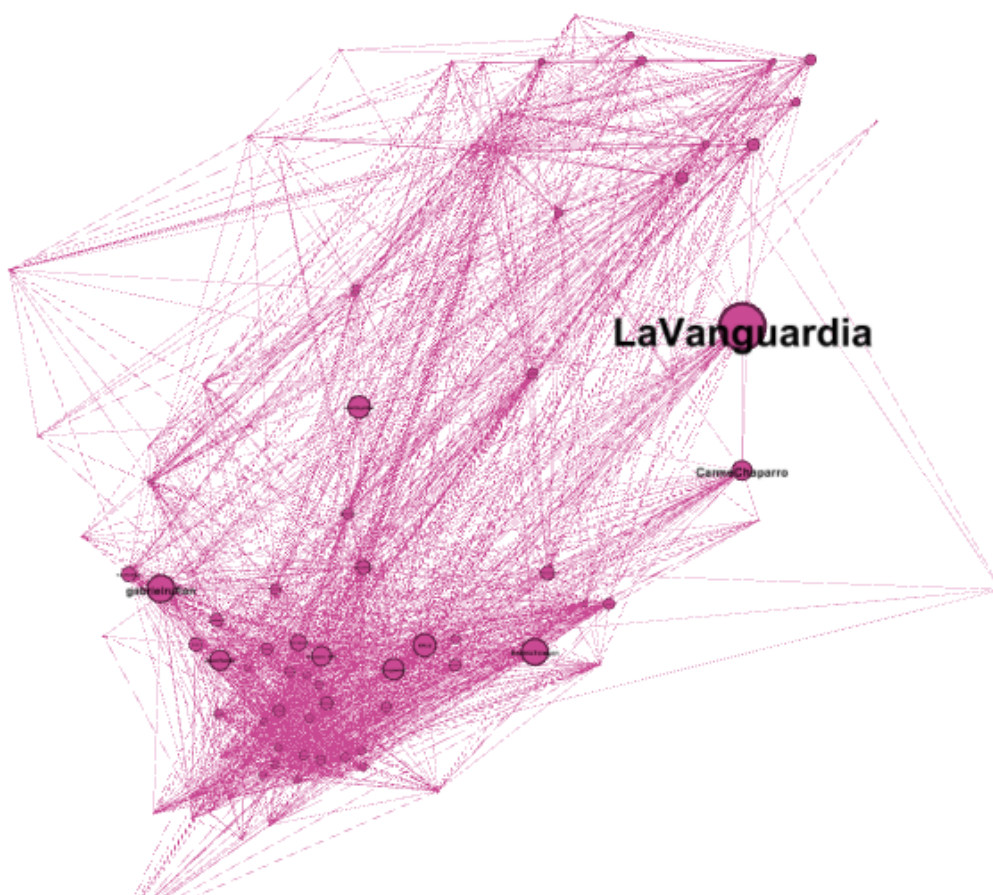
De todos modos, es posible obtener otro tipo de resultados en función de la resolución aplicada a la hora de calcular las comunidades que conforman la red política española. Por lo tanto, además por cómo trabaja Gephi con la modularidad, este resultado se debe al grado de granularidad establecido.

En la anterior red obtenida (Figura 1) la resolución aplicada fue de 1.0. Si se experimenta con este parámetro nos podemos encontrar con algún que otro resultado digno de mención.

En este caso, si se aplica una resolución de 0.5 se puede obtener una subred de partidos independentistas separados de la antigua red de Unidas Podemos.

En la Figura 49 se puede observar la subred de la red inicial que forman estos partidos independentistas.

Figura 49. Red inicial de partidos independentistas



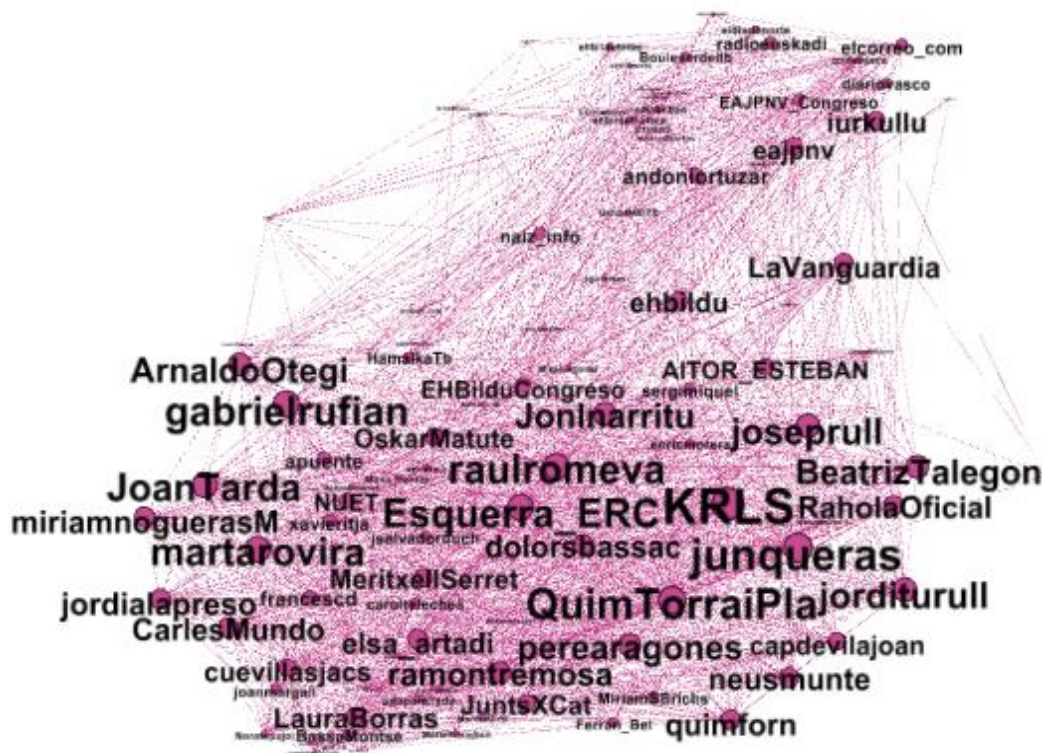
Fuente: Elaboración propia

Como podemos observar, salvo el periódico ‘La Vanguardia’, ningún usuario goza de gran prestigio en la red global de la política española si tomamos como referencia el número de seguidores de la red.

Si se da forma al grafo, de tal forma que observemos qué usuarios son los más relevantes en función de los seguidores en esta subred, se puede obtener otro tipo de resultado.

En la Figura 50 se observa a la subred de partidos independentistas catalanes y vascos en función del número de seguidores en la subred.

Figura 50. Red de partidos independentistas en función del número de seguidores



Fuente: Elaboración propia

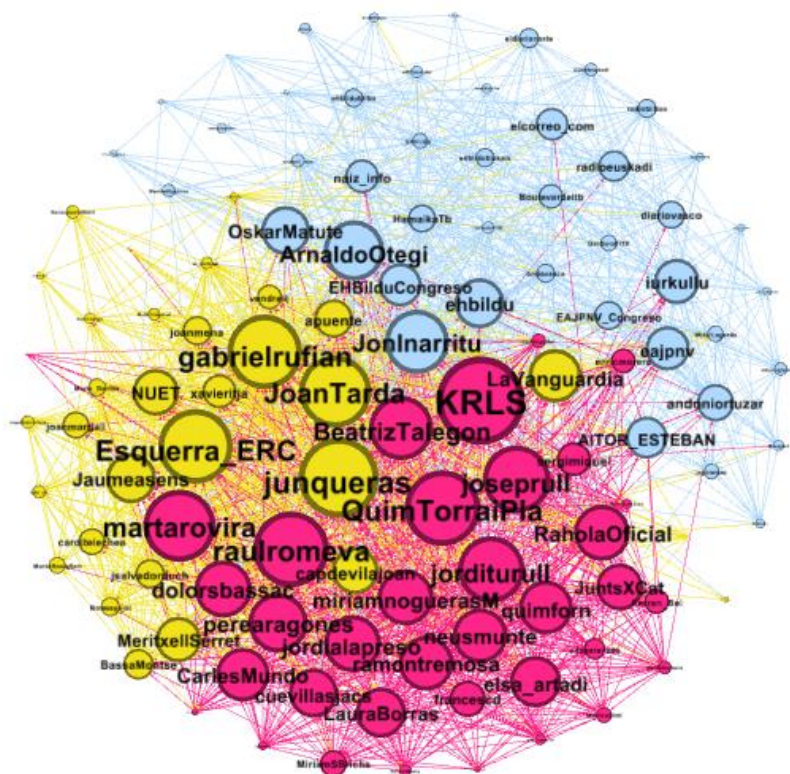
Podemos apreciar como la mayoría de los usuarios adquieren prestigio dentro de su propia comunidad, dada la sólida estructura de seguidos entre partidos independentistas.

Sin embargo, el periódico 'La Vanguardia' pierde protagonismo, es decir, sirve como punto de conexión entre partidos independentistas y el resto de la red. En este caso, obtendría un alto grado de centralidad de intermediación, es decir, gran parte de los usuarios de la subred dependen de su actuación para mantener una relación con actores externos.

Se puede conseguir un resultado más exacto probando con la resolución de detección de comunidades.

En la Figura 51 podemos ver como se distinguen tres grandes grupos: Esquerra Republicana de Catalunya, Junts per Catalunya y los partidos políticos vascos.

Figura 51. Partidos independentistas en función de la modularidad



Fuente: Elaboración propia

14.6. Resumen

Para concluir el análisis estático de la red, podemos enumerar una serie de hechos notables. Es importante la estructura que presenta la red, tanto global como por partidos políticos. Los actores más externos de la red son, en su mayoría, cuentas institucionales de los diferentes partidos, las cuales no suelen mantener relación con usuarios externos a la subred de su partido político. Sin embargo, los usuarios que suelen hacer de puente entre el partido y el resto son los medios de comunicación y periodistas. Por otro lado, los líderes y representantes de cada partido hacen de hilo conductor entre las cuentas institucionales y los medios y periodistas.

Por lo tanto, se puede concluir, que las cuentas oficiales de partido se encuentran, en su gran mayoría en las zonas más periféricas de la red, dejando la zona central a aquellas cuentas que consiguen la comunicación entre las diferentes fuerzas. De todas formas, no todas las subredes siguen un mismo posicionamiento en la red. Todas tienen similitudes con el resto, pero también diferencias.

En el caso de la red de Unidas Podemos, destaca por encima de todos la figura de Pablo Iglesias, incluso por delante de la cuenta oficial del partido. El siguiente escalón lo conforman representantes del partido, el medio 'El Diario' y el periodista Jordi Évole. El resto de cuentas tienen una relevancia un tanto anecdótica, lo que nos indica una estructura muy similar a la que tiene realmente el partido como estructura jerárquica. Es un partido que apuesta por el consenso y el diálogo con sus simpatizantes, pero no dejan de existir rangos dentro del partido.

Por otro lado, el PSOE muestra algunas similitudes como la imagen de Pedro Sánchez como la más relevante, pero existe mayor desigualdad. De las siguientes cinco cuentas más destacadas, ninguna es un representante o la cuenta del partido en alguna ciudad. Destacan medios como El País o La SER y la cuenta oficial del partido.

La comunidad del PP muestra una red menos desigual, pero es peculiar que muchos de los actores más destacados ya no estén expuestos a un primer plano en la política española como, por ejemplo, Esperanza Aguirre o Cristina Cifuentes.

Sin embargo, con Ciudadanos encontramos bastante disparidad. Sin duda es la red con menor desigualdad. Ninguna cuenta sobresale en exceso por encima del resto. Despuntan ligeramente Albert Rivera, Inés Arrimadas y la cuenta oficial del partido, pero muy cerca aparecen las cuentas de los representantes con mayor peso en el partido, como Girauta.

Por último, VOX presenta una peculiaridad en común con Unidas Podemos, Ambos partidos poseen cuentas de ciudades más pequeñas que el resto y, en principio, menos relevantes, pero con mayor relevancia en la red. Estos dos partidos siguen una política distinta a la común en los partidos tradicionales como PP o PSOE que albergan su imagen en un menor número de integrantes. Algunos ejemplos son la cuenta de Vox en Colmenar Viejo (49.000 habitantes) o la cuenta de Unidas Podemos en Lluçmajor (36.000 habitantes). Otro punto en común entre Unidas Podemos y VOX, y fruto de la posición ideológica de ambos, son las fuerzas con mayor número de 'influencers' en su red. Esto

puede ser debido a la facilidad de ciertas cuentas para posicionarse en estos partidos. Es, en principio, más complicado ser defensor de otro partido político debido a los límites políticos más difusos que marcan. Sin embargo, tanto con Unidas Podemos como con VOX, su posición ideológica es bastante evidente.

En el caso de PSOE y PP siguen una estructura muy semejante. Ambos albergan grandes medios de comunicación y existen grandes diferencias de prestigio entre sus integrantes. Es decir, las cuentas menos seguidas, normalmente las cuentas institucionales, tienen mucho menos peso que el resto de cuentas del partido, incluyendo a los líderes de estos. Como detalle, es curiosa la aparición de Coalición Canaria en la red del PSOE, ya que es un partido que no suele posicionarse demasiado sobre su ideología y no suele entrar en comparación con el resto de partidos.

En el caso de Ciudadanos, encontramos la contraposición con PSOE Y PP. Es, claramente, la comunidad en la que menor desigualdad existe entre sus usuarios en cuestión de prestigio. La gran mayoría de los usuarios poseen una cantidad de seguidores muy parecida dentro de su red, lo que refuerza la idea de que es la red con mayor cohesión, es decir, la mejor comunicada entre sus integrantes.

Por último, aunque no al nivel de PSOE y PP, VOX sorprende con la cantidad de medios de comunicación, periodistas, programas de radio y organizaciones en su red. Todas estas llevan a cabo la misma función dentro de la red de VOX; comunicar al resto de la red con el partido.

15. Análisis dinámico de la red

Una vez analizada estáticamente la red, se da paso a la fase de análisis dinámico, es decir, el análisis de grafos representantes de eventos de carácter político.

Como se ha descrito en el apartado 8, el análisis consta de 3 grafos. El primero proyecta la imagen de la red política española durante el debate de Radio Televisión Española, el segundo, de igual forma, durante el debate de Atresmedia y, por último, el tercero presenta la conducta de la red durante el día de las elecciones.

Se comenzará elaborando un análisis de cada uno de los eventos para, posteriormente, entrar en comparación con el resto.

15.1. Debate electoral en RTVE el 22 de abril de 2019

El primer suceso es el debate de Radio Televisión Española, celebrado el día 22 de abril a las 22 horas. La representación de la red se puede apreciar en la Figura 55.

De los 1688 integrantes de la red, solo participaron durante el evento 737, lo que representa un 43,66% del total. Se han detectado un total de 6265 interacciones entre los integrantes de la red. De ellas, 1875 únicas, es decir, cuyo origen y destino no es el mismo.

En primera instancia, se pueden detectar algunas comunidades de forma clara. La más grande, representado un 20,08% de la red, es la del Partido Popular. En segunda posición, se encuentra la comunidad de Unidas Podemos con un 17,23%. Muy próximas, se hallan las comunidades de Ciudadanos y PSOE con 16,69% y 16,55% respectivamente. Alejada del resto, se observa a la de VOX con un 7,46% del total.

Como se puede ver, hasta ahora solo aparecen comunidades referentes a los cinco partidos políticos más influyentes en este momento. Si aunamos porcentajes, estos partidos representan un 78% de la red, más de tres cuartas partes.

Ahora se puede centrar el foco en qué reside el 22% restante. Si se continúa desglosando por comunidades, la siguiente sería la formada por medios digitales, influencers y comunicadores de izquierdas con un 3,39%. En ella participan @martaflich, economista española que participa con el diario Huffington Post (@ElHuffPost), perteneciente a esta comunidad, realizando videoblogs sobre la situación económica en España. Otro actor destacado es el 'influencer' @gerardotc, guionista del programa de televisión 'Late Motiv' y escritor en la revista digital @ctxt_es, integrante de la comunidad. Destaca el influencer @Xuxipc y sorprende la aparición de @RitaBosaho, diputada por Unidas Podemos, ya que es la única representación política en la comunidad.

A continuación, aparece la comunidad formada por integrantes de EHBildu, JuntsXCat ERC con un 3,12%. Destaca la cuenta oficial de EHBildu (@ehbildu), el diputado de ERC, @gabrielrufian y el presidente de la Generalitat, @QuimTorraPla.

Seguidamente, con un 2,44%, se encuentra la comunidad formada por las distintas cuentas de RTVE y la organización contra la desinformación con @maldita_es a la cabeza. En esta subred destaca, sobre todo y como es lógico, la primera sección comentada.

La siguiente comunidad es la conformada por cuentas de Atresmedia como @A3Noticias o @DebateAlRojoVivo, y la productora de televisión y contenido periodístico, Newtral. Esta productora es dirigida por la periodista Ana Pastor (@anapastor), que también pertenece a esta comunidad.

Por último, aparecen cuatro pequeñas comunidades que representan a Compromís, EnMarea, EAJPNV y Coalición Canaria con un 1,36%, 1,36%, 1,09% y 0,81% respectivamente.

El porcentaje restante lo conformarían actores aislados como, por ejemplo, @IgnatiusFarray que solo interactúa con el programa de radio donde colabora @vidamoderna.

Si se centra el foco en los cinco partidos más importantes, que en este caso coincide con las cinco comunidades más grandes, se pueden detectar los actores principales en la Tabla 1 y los tres con mayor valoración por cada parámetro medido en la Figura 54.

Durante el debate, el grado de interacción que reciben los distintos usuarios de la red indica el número de actores que han interactuado con el usuario en cuestión y no el número de interacciones totales que recibieron. Es destacable que las tres primeras cuentas sean las oficiales de partido, incluso la cuarta es @ahorapodemos, la cuenta oficial de Unidas Podemos.

También es llamativo que los siguientes actores sean los cuatro líderes de estos partidos, exceptuando a la cuenta de 'El País', que aparece en la séptima posición entre Pedro Sánchez y Albert Rivera. Pablo Casado ocupa el quinto puesto y Pablo Iglesias el noveno. Sin embargo, VOX no aparece hasta la undécima posición mientras que, Santiago Abascal desciende hasta la posición 52. Esto se explica mejor si se tiene en cuenta que VOX no pudo participar en el debate por falta de representación parlamentaria.

También se ha tenido en cuenta el grado de salida de cada usuario, es decir, el número de interacciones creadas por el usuario con distintos usuarios de la red. A este parámetro lo hemos llamado grado de actividad. Si un usuario tiene 10 interacciones, pero solo con dos usuarios distintos, el grado de salida será dos. En general, entre PSOE, PP y Unidas Podemos se intercalan las posiciones en este aspecto. Es curioso que Ciudadanos no se introduzca en la lista hasta la posición 26, de hecho, entre los cincuenta primeros usuarios,

solo obtiene representación con uno. Esto muestra la poca interacción de la comunidad de Ciudadanos en durante el debate, llegando a marcar registros similares a los de VOX.

Si se toma como referencia la medida de centralidad de intermediación, no encontramos un representante del PSOE hasta la decimotercera posición con la cuenta oficial del partido. Sorprende que VOX se introduzca entre los 25 mejor posicionados con @vox_es (posición 22). Pero sobre todo es extraño que Ciudadanos no muestre representación hasta la posición 84. Esto es entendible si recordamos que la comunidad de Ciudadanos es la más cohesionada y en la que menor número de medios aparecen. Como ya se ha comentado, los medios de comunicación suelen ejercer de puente con otras comunidades, por lo tanto es comprensible este hecho.

Entre los veinticinco primeros integrantes, doce son de Unidas Podemos, diez del Partido Popular, dos del Partido Socialista y uno de VOX. El posicionamiento estratégico de los dos primeros es dominador. Sin embargo, como se verá a continuación, no por ello los demás partidos pierden protagonismo.

Hasta ahora se ha visto el grado de centralidad o posicionamiento en la red de los distintos integrantes de esta. En el caso del prestigio en la red, la situación cambia.

Si tomamos como referencia de grado de prestigio el parámetro Pagerank, hasta la sexta posición no se encuentra representación de Unidas Podemos, en este caso, con @ahorapodemos. Por otro lado, el primer representante de VOX alcanza la duodécima posición con @vox_es. Entre los veinticinco primeros actores se puede denotar alguna diferencia. Unidas Podemos es el partido con mayor representación, en este caso nueve, seguido de PSOE con cinco, Partido Popular y VOX con cuatro y Ciudadanos con tres.

Sin embargo, si se toma un listado intermedio y se ciñe a los diez primeros, habría triple empate entre Ciudadanos, PSOE y PP con tres mientras que Unidas Podemos que, a priori parecía vencedor, queda tan solo con uno.

Tabla 1. Clasificación de actores más importantes por subred en función de los parámetros evaluados durante el debate de RTVE

Subred	Usuarios más importantes	Interacción	Actividad	Centralidad de intermediación	Pagerank
PSOE	PSOE	3º (100)	4	684	3º (0,022)
	AlejandroSolerM	0	2º (27)	0	0,0029
	Pmorla68	0	3º (20)	0	0,0026
PP	populares	1º (118)	3	3º (3969)	2º (0,027)
	pablocasado_	69	4	2º (4229)	0,019
	CarmenRiolobos	1	1º (27)	0	0,000269
Ciudadanos	Ciudadanos	2º (107)	2	395	1º (0,029)
	Albert_Rivera	48	0	0	0,019
	Cs_Andalucia	3	2	0,5	0,0005
Unidas Podemos	ahorapodemos	76	14	1º (6681)	0,014
	Pablo_Iglesias_	37	1	793	0,005
	podem_cat	22	5	897	0,0027
VOX	vox_es	27	2	99	0,0062
	Santi_ABASCAL	6	2	94	0,003
	madrid_vox	11	3	56	0,0018

Fuente: Elaboración propia

Figura 54. Los tres actores con mayor valoración en los cuatro parámetros durante el debate de RTVE



Fuente: Elaboración propia

Durante el debate se trataron cuatro temas:

- Política económica, fiscal y empleo
- Política social, estado del bienestar, pensiones e igualdad
- Política territorial
- Regeneración democrática y pactos postelectorales

A lo largo del debate, se produjeron un total de 559.000 tweets y participaron 131.000 usuarios. De este total, en nuestra red se han recogido 6265 tweets de 737 usuarios diferentes.

El volumen de tweets no fue lineal y alcanzó su nivel máximo a las 23:00 seguido de las 23:30.

El primero coincide con un tema que provoca uno de los momentos más tensos del debate, la supuesta reunión del PSOE con los partidos independentistas. En ese momento, tanto Pablo Casado como Albert Rivera mantienen una postura agresiva y su objetivo en el partido socialista, en este caso representado por Pedro Sánchez. Por el contrario, Pablo Iglesias llama a la cordura y al respeto por la educación ante el espectador.

El segundo momento coincide con el minuto final para cada líder y la conclusión del debate. Muy comentado ha sido el minuto de Albert Rivera, con algún que otro eslogan propio del partido con un mensaje muy enfático de cara al espectador.

Por lo tanto, a lo largo del desarrollo del debate destacó, por encima del resto, el conflicto catalán.

Durante el resto del debate, Ciudadanos mantuvo su postura activa y ofensiva hacia el resto de partidos, sobre todo PSOE. Por otro lado, PP realizó algún que otro intento, pero se vio superado por el partido de Albert Rivera. Por el contrario, PSOE se limitó a no entrar demasiado en polémica y salió airoso de los temas más candentes. Unidas Podemos optó por una posición apaciguadora, tratando de dar una imagen muy diferente a la mostrada en anteriores debates, poniendo en algún apuro a Pedro Sánchez, buscando así llamar la atención del votante socialista.

Con los datos recogidos se puede afirmar que Ciudadanos engloba los usuarios más prestigiosos durante el debate, pero acumulando gran parte de todo su prestigio en estos, ya que en niveles de interacción alcanza niveles tan bajos como los de VOX, que no participaba en el debate. Esto refleja agresividad con la que se desenvuelve Albert Rivera durante el debate, es decir, pocos usuarios, pero se hacen notar. Es la situación contraria

a Unidas Podemos, que acumula una gran cantidad de usuarios con alto nivel de interacción pero que carecen de cierto prestigio.

Existe disparidad de opiniones sobre quién salió vencedor del debate. El periódico El País da como ganador a Unidas Podemos seguido de Ciudadanos. Por el contrario, El Mundo y La Vanguardia dan como vencedor a Albert Rivera seguido de Pablo Casado en el caso del primero y Pablo Iglesias del segundo. También hay medios que no acaban de apostar por un candidato como, por ejemplo, El Periódico, que ve un empate entre Ciudadanos y Unidas Podemos.

Como se ve, el vencedor del debate depende del medio. El bloque conservador apuesta por la victoria de Albert Rivera, incluso por el buen papel de Pablo Casado, sin duda el gran perdedor del evento por la mayoría de los expertos. Sin embargo, los medios más progresistas no ven tan clara la victoria de Albert Rivera, al que su “descarada actuación durante el debate acabó perjudicándole”, según algunos periodistas.

Si los expertos en la materia dudan, no podemos declarar un ganador de forma evidente, pero sí que destacaron dos de ellos, Albert Rivera y Pablo Iglesias, por diferentes motivos, ya que plantearon una estrategia totalmente diferente.

También podemos entender mejor el posicionamiento de ciertos medios en su correspondiente comunidad. Como se ha comentado, tanto La Vanguardia, posicionada en la red del PP, como El Mundo, posicionada en la de VOX, apuestan por Rivera como ganador. También es más entendible que El Mundo apueste por un Pablo Casado sólido durante el debate.

Por otro lado, El País y El Diario discrepan y ven a Pablo Iglesias como ganador y no dudan de que Pedro Sánchez salió airoso, ya que “tenía mucho que perder”. Tampoco niegan el buen papel de Albert Rivera, posicionado en segundo lugar.

Esto nos hace indicar que el posicionamiento de gran parte de los medios es por afinidad y no por grado de interacción durante el debate, al menos en estos casos.

Por último, dada la poca interacción de la red de Ciudadanos, parece curioso que acabase proclamándose ganador por la mayoría de los expertos. Parece más lógico que, partidos como PSOE o PP que mantienen niveles altos de prestigio y posicionamiento en la red, lo hubiesen reflejado en el resultado final. Por lo tanto, la relevancia en nuestra red en este caso no es representativa de la realidad.

15.2. Debate electoral en Atresmedia el 23 de abril de 2019

El segundo evento es el debate de Atresmedia, llevado a cabo el día 23 de abril a las 22 horas. La representación de la red se puede apreciar en la Figura 57.

De los 1688 integrantes de la red, solo participaron durante el evento 767, lo que representa un 45,43% del total. Se han detectado un total de 6190 interacciones entre los integrantes de la red. De ellas, 2118 no repetidas.

En primera instancia, se pueden detectar algunas comunidades de forma clara. La más grande, representando un 19,82% de la red, es la del Partido Popular. En segunda posición se encuentra la comunidad de Unidas Podemos con un 19,43%. Cercanas a estas, se hallan las comunidades de PSOE y Ciudadanos con 18,51% y 17,99% respectivamente. Como en el anterior debate, se observa a lo lejos a VOX con un 9%.

Hasta ahora solo aparecen comunidades referentes a los cinco partidos políticos más influyentes en este momento. Si sumamos todos los porcentajes, estos partidos acumulan un 84,75 % de la red.

Si se centra la mira en el 15,25 % restante y se continúa desglosando por comunidades, la siguiente, con un 2,48%, sería la formada por los partidos independentistas EHBildu y JuntsxCat donde los actores más destacados son @ArnaldoOtegi, @JonInarritu, @OskarMatute, todos de EHBildu. La aparición de JuntsXCat es meramente anecdótica.

La siguiente comunidad, con un 2,22%, es la conformada por medios digitales, influencers y comunicadores de izquierdas. Como en el debate de RTVE, los actores más destacados en esta comunidad son @martaflich, @ElHuffPost, @gerardotc y el periodista @AntonioMaestre.

Seguidamente, vuelve a aparecer una comunidad en común con el debate de RTVE. Con un 1,56% de la red, aparece la conformada por cuentas de Atresmedia como @A3Noticias o @DebateAlRojoVivo, y la productora de televisión y contenido periodístico, Newtral, junto con su directora, la periodista Ana Pastor (@anapastor), que también pertenece a esta comunidad. Sorprende la aparición del influencer @FrayJosepho, conocido por su ideología de derechas o incluso más por su oposición a la de izquierdas. De hecho, en el debate de RTVE aparece situado en la comunidad de VOX.

Más aletargados aparecen las comunidades de EnMarea, Compromís y EAJPNV ocupando un 1,04%, 0,65% y 0,39% de la red, respectivamente.

Sorprende la disgregación de ERC en este evento, en el que actores principales del partido como @junqueras o @gabrielrufian aparecen separados en comunidades de muy poco peso, por debajo incluso de 0,5% de la red.

Si se apunta a los cinco partidos con más poder del momento se pueden extraer algunos datos, se pueden detectar los actores principales en la Tabla 2 y los tres con mayor valoración por cada parámetro medido en la Figura 56.

Durante el debate de Atresmedia, respecto al grado de interacción, las primeras cuentas son las oficiales de los partidos, ya que en cuarta posición se encuentra @PSOE con 88 interacciones. Sin embargo, es más notable la diferencia en cuanto a los líderes de partido. En este caso, las tres siguientes cuentas son las de Pablo Casado, Albert Rivera y Pedro Sánchez con 60, 60 y 43 interacciones respectivamente. Esto concuerda con lo visto en el anterior debate, pero llama la atención que la cuenta de Pablo Iglesias no aparezca hasta la posición 57 con siete interacciones, mientras que en el anterior aparecía entre los diez primeros.

Este último dato es más significativo si se observa que Santiago Abascal empató en interacciones con Pablo Iglesias, aun no participando en el debate. Por otro lado, la cuenta oficial de VOX no aparece hasta la octava posición, tres posiciones por encima que el día anterior.

En función del grado de actividad, sin duda, el Partido Popular posee a los usuarios más activos durante el debate, sumando trece representaciones entre las veinte primeras. Las siete restantes se las reparten PSOE con cinco y Ciudadanos y Unidas Podemos con una. En el caso de VOX, vuelve a la normalidad al no participar durante el debate, ya que muestra representantes hasta la posición 92.

Respecto al grado de centralidad y tomando como referencia la centralidad de intermediación, el Partido Socialista domina este segmento ya que ocupa las cinco primeras posiciones. El Partido Popular no aparece hasta la séptima posición con la cuenta oficial del partido y Unidas Podemos en la decimoquinta posición con la cuenta oficial de Izquierda Unida, mientras que Ciudadanos no muestra representación hasta la decimonovena posición con la cuenta del partido en Andalucía. En cualquier caso, todos muy por debajo de las 5 primeras cuentas.

Esto puede ser un reflejo de las diferencias en cuanto a la estructura jerárquica dentro de cada partido. Como ya se ha comentado en apartados anteriores, en el caso de Ciudadanos no es necesaria la figura de un líder para comunicar al resto de los usuarios, mientras que en el PSOE vemos como si es indispensable la imagen de un usuario o grupo de usuarios que permita la comunicación entre ellos. Por lo tanto, parece normal que existan usuarios con un alto grado de centralidad de intermediación en la comunidad del PSOE. En este evento, es el único partido al que le sucede. A parte de los tres actores mencionados, Cristina Narbona y Francisco Polo también son actores clave dentro de la red.

Si hablamos de prestigio en la red en función del Pagerank, la situación es diferente. El reparto de prestigio por partidos está bastante igualado, aunque Ciudadanos domina con holgura al tener tres representaciones con puntuaciones sobresalientes entre los cinco

primeros, con las cuentas oficiales del partido en España y Andalucía y la de Albert Rivera.

Al igual que con el grado de interacción, Pablo Iglesias queda muy relegado, en concreto hasta la posición 80, muy por detrás de Santiago Abascal (18), la cuenta oficial de VOX (12) o el resto de líderes, todos entre los diez primeros. De esta forma acaba asumiendo un rol secundario a lo largo del debate, algo que puede ser entendible dada su postura conciliadora durante el debate. Esto último también puede reflejar que Pablo Iglesias interactúe con usuarios en la red, a priori, menos relevantes. Si a esto se le suma la escasa interacción con otros actores, origina un reducido prestigio en la red.

Tabla 2. Clasificación de actores más importantes por subred en función de los parámetros evaluados durante el debate de Atresmedia

Subred	Usuarios más importantes	Interacción	Actividad	Centralidad de intermediación	Pagerank
PSOE	PSOE	88	7	1º (6378)	0,018
	socialistes_cat	15	7	2º (4313)	0,0016
	miqueliceta	1	3º (26)	3º (3753)	0,00026
PP	populares	1º (121)	5	1260	2º (0,036)
	pablocasado_	60	2	47	0,013
	PPopularAvila	6	2º (29)	591	0,0006
Ciudadanos	Ciudadanos	3º (99)	2	570	1º (0,045)
	Albert_Rivera	60	1	77	3º (0,031)
	Cs_Andalucia	23	4	574	0,021
Unidas Podemos	ahorapodemos	2º (114)	2	252	0,024
	AdaColau	17	1	35	0,0021
	Podem_	4	2	1	0,01
VOX	vox_es	39	2	392	0,008
	Santi_ABASCAL	7	5	426	0,004
	voxnoticias_es	7	6	283	0,0025

Fuente: Elaboración propia

Figura 56. Los tres actores con mayor valoración en los cuatro parámetros durante el debate de Atresmedia



Fuente: Elaboración propia

Durante el debate se trataron tres apartados:

- Programa electoral
- Modelo territorial
- Pactos

A lo largo del debate, se produjeron un total de 653.000 tweets y participaron 145.000 usuarios. De este total, en nuestra red se han recogido 6190 tweets de 767 usuarios diferentes.

El volumen de tweets no fue lineal y alcanzó su nivel máximo a las 22:20 seguido de las 22:45.

El primero coincide con otro momento tenso entre Pedro Sánchez y Albert Rivera coincidiendo con el apartado de empleo, aunque en ese momento se hable de todo menos de esto. De hecho, al igual que en el anterior debate, vuelve a salir a debate el conflicto catalán y con los mismos protagonistas.

El segundo momento corresponde con un intercambio de opiniones entre Pablo Iglesias y Pedro Sánchez sobre impuestos, aunque acaba siendo una disputa entre el primero y Albert Rivera sobre las continuas interrupciones de ambos.

En el caso de Ciudadanos se vuelve a repetir el mismo patrón, es decir, poca interacción pero con gran impacto, ya que acumula a los usuarios con mayor prestigio. Por otro lado, Unidas Podemos, no solo no obtiene un alto grado de prestigio, sino que su nivel de interacción se reduce a niveles cercanos a VOX.

En el caso de PSOE y PP, sus integrantes mantienen un nivel alto de interacción y albergan usuarios con cierto prestigio. Es entendible si observamos que ambos acogen medios de comunicación muy activos durante el debate; El País y El Diario en el caso de PSOE y Antena3 en el de PP.

Esto nos puede dar una pista del motivo del nivel de interacción de Unidas Podemos en el debate de RTVE, en el que El Diario se posiciona en la red del partido de Pablo Iglesias. Es evidente que el hecho de englobar a un medio influyente en la red hace aumentar los niveles de interacción y mostrar resultados, más si cabe, sujetos a interpretación.

En cuanto a VOX, sigue manteniéndose al margen, aunque tanto su nivel de interacción y de prestigio muestran un leve subida. Esto se puede deber a que, en forma de protesta y para desviar la atención del debate de Atresmedia, VOX celebra un mitin en Las Rozas en las que acaba reuniendo más de cinco mil personas. Durante el evento, los actores más activos y que más crecimiento alcanzaron respecto al día anterior fueron la cuenta oficial de VOX y Santiago Abascal. Los temas tratados fueron distintos a los del debate de RTVE. El 22 de abril, el partido se limitó a echar en cara a RTVE el no poder participar en el debate y a utilizar un discurso desafiante contra sus oponentes, sobre todo con PSOE, al que se le ofrece un cara a cara públicamente. Aun así, el partido de Santiago Abascal continúa muy lejos de la irrupción que se producirá el día de las elecciones.

También se ha mencionado que el Partido Popular posee, generalmente, a los usuarios más activos. Estos usuarios muestran dos comportamientos que se repiten durante todo el debate. El primero es el apoyo incondicional hacia Pablo Casado, al que autoproclaman como claro vencedor del evento. El segundo es comentar negativamente todo lo que plantea el PSOE en el debate, en especial hacia la figura de Pedro Sánchez. Por otro lado, hay poca interacción con Ciudadanos, algo que parece favorecer lo que sucederá el día de las elecciones, es decir, el paso de una gran cantidad de votantes del Partido Popular a Ciudadanos. El posible votante de PP no ve como amenaza al partido de Albert Rivera, sino como una alternativa.

Al igual que en el debate de RTVE, el vencedor vuelve a tener muchos nombres. En este caso, se une a la pelea Pablo Casado. Prácticamente, ningún medio duda de que Pedro Sánchez salió perjudicado del evento celebrado en Atresmedia. Sin embargo, medios como El Mundo, Onda Cero y ABC dan como vencedor a Rivera. Los dos primero afirman una remontada de Pablo Casado, mientras que ABC solo destaca al líder de Ciudadanos. OkDiario manifiesta una clara victoria de Pablo Casado y una gran caída de la izquierda. Por último, El País y una encuesta recogida por La Vanguardia que engloba a los distintos medios en España dan como vencedor a Pablo Iglesias, seguido de Pablo Casado.

15.3. 28 de abril - Día de elecciones

El tercer y último evento es el día de las elecciones, llevadas a cabo el día 28 de abril. La representación de la red se puede apreciar en la Figura 59.

De los 1688 integrantes de la red, solo participaron durante el evento 827, lo que representa un 48,99% del total. Aunque fue el día que más tweets se publicaron en Twitter desde la campaña electoral, en nuestra red solo se han detectado un total de 4386 interacciones entre los integrantes de la red. De ellas, 1787 no repetidas.

Se pueden observar notoriamente ciertas comunidades. La mayor, proyectando un 18,38% de la red, es la de Ciudadanos. En segunda posición se encuentra la comunidad del Partido Popular con un 15,11%. Las dos siguientes son PSOE y Unidas Podemos con 13,42% y 11,73% respectivamente.

A diferencia de los debates, la siguiente comunidad no es la de VOX, sino la formada por la escisión de Unidas Podemos en la que aparece Izquierda Unida, medios de comunicación y periodistas afines a la ideología de izquierdas como @eljueves y @AntonioMaestre, influencers de clara tendencia de derechas como @UnTioNormal_XD y de izquierdas como @Panik81. También la conforman actores como Anabel Alonso y Carlos Bardem afines a partidos de izquierdas. Esta subred abarca un 9,92% del total de la red.

Si se obvia esta red intermedia y se cuenta con la de VOX, todas suman un 68%, sensiblemente menos que en los debates. Además de la red situada entre Unidas Podemos y VOX, se pueden encontrar otras comunidades relevantes en el 22% restante.

Muy cercana a la anterior, aparecería la comunidad de VOX con un 9,31%. Las siguientes comunidades serían la de EHBildu con un 2,78%, la formada por medios de Atresmedia y Newtral con un 2,42%, la formada por el programa televisivo 'Todo es mentira', la revista digital JotDown en España y el periódico 'El Mundo'.

A continuación, aparecen JuntsXCat con un 2,06%, ERC con un 1,45% y Compromís, Coalición Canaria, EAJPNV y la organización anti bulos, todas con 0,97%.

Por último, se observa la comunidad de EnMarea con un 0,73% de la red.

Si se apunta a los cinco partidos con más poder del momento se pueden extraer algunos datos, se pueden detectar los actores principales en la Tabla 3 y los tres con mayor valoración por cada parámetro medido en la Figura 58.

En contraposición a lo visto en los debates, en este caso no se denota claramente que las cuentas oficiales de los partidos sean las que mayor grado de interacción tengan. De hecho, aparecen cuentas autonómicas como la de Ciudadanos en Andalucía por delante de la del PSOE. Esto tiene una sencilla explicación, se ha comentado que el grado de interacción es el número de usuario distintos que interaccionan con otro. Por lo tanto, esto

nos da una pista de la tipología de interacción que reciben PSOE y Ciudadanos. Parece que la cuenta oficial del partido socialista recibe interacciones de un menor número de usuarios que la de algunas cuentas autonómicas, pero esto no impide que, seguramente, @PSOE reciba un mayor volumen de interacciones al ser la cuenta más visible del partido junto con la del líder. Esto nos indica y refuerza lo ya comentado, Ciudadanos posee la red más cohesionada, donde prácticamente todos los usuarios que la conforman se siguen entre sí y existe poca desigualdad en su red. Por lo tanto, parece lógico que establezca interacción con un mayor número de usuarios que las cuentas posicionadas en la comunidad del PSOE, donde destacan solo ciertas figuras como los principales representantes de partido y algún medio de comunicación.

Al contrario que en los debates y a pesar de la aparición de una comunidad no correspondiente con un partido antes que VOX, el partido liderado por Santiago Abascal escala posiciones y alcanza un nivel similar al del resto de partidos en cuanto a nivel de interacciones. Una prueba de ello es el posicionamiento de @vox_es en la cuarta posición y @Santi_ABASCAL entre las quince primeras. Para explicar mejor el contexto, la cuenta de Santiago Abascal está por delante de la de Pablo Iglesias, que no aparece hasta la posición 29. Incluso cuentas de representantes autonómicos como las de Inés Arrimadas (12) o Ignacio Aguado (22), aparecen antes que la del líder de Unidas Podemos.

Se puede evidenciar el contraste entre Ciudadanos y Unidas Podemos, donde los primeros obtienen una gran cantidad de interacciones de distintos usuarios de la red, mientras que los segundos, a pesar de tener seguramente gran cantidad de interacciones a través de la plataforma, no consiguen interaccionar con cierta pluralidad en la red política española.

En función del grado de actividad, sin duda, Ciudadanos, al igual que el Partido Popular en el debate de Atresmedia, posee a los usuarios más activos, sumando diez integrantes entre los veinte primeros. Los diez restantes se los reparten la cuarta comunidad en tamaño con cuatro, PSOE con dos, PP con dos y Unidas Podemos y VOX con una.

Si se toma como referencia la medida de centralidad de intermediación, aunque predomina la presencia de PSOE Y Ciudadanos, sorprende la aparición de la organización anti bulos con tres cuentas entre las siete primeras. Esta organización lo que pretende es desmentir aquellas informaciones falsas malintencionadas sobre la política con la intención de hacer política de una forma más justa. Comúnmente, en las vísperas de elecciones, aparece alguna que otra noticia falsa, independientemente del partido, con la intención de tergiversar el discurso de un determinado partido. Por lo tanto, parece normal que este tipo de usuarios cuenten con una posición privilegiada en un día tan señalado como el de las elecciones y sirvan como nexo entre distintos usuarios en la red.

Otra disparidad respecto a los debates es que, al igual que el PSOE poseía usuarios posicionados con cierto privilegio, en este caso, sorprende que le ocurra lo mismo a Ciudadanos. Esto podría evidenciar la avanzada de última hora del partido en días previos

a las elecciones, tratando de interactuar más con usuarios no integrantes de su partido. Todo lo contrario sucede con el Partido Popular y su líder Pablo Casado, que en este aspecto desaparece por completo. Le superan partidos como Unidas Podemos o incluso VOX, que poseen un usuario mejor posicionado que el resto de integrantes del Partido Popular.

En cuanto al prestigio de la red en función del Pagerank, el reparto de prestigio aparece bastante equilibrado entre las distintas fuerzas, aunque sorprenden dos aspectos. El primero es la gran representación de VOX, acumulando cinco usuarios entre los veinte primeros. Es cierto que, durante el día de las elecciones, se ha posicionado a la cadena RTVE en la comunidad de VOX por su relación de interacción y dos de las cinco mencionadas están asociadas a la cadena (@rtve y 24h_tve). Repasando, a través de búsquedas avanzadas en Twitter, no se observa ningún posicionamiento claro de RTVE respecto a VOX durante el 28 de abril. Por lo tanto, parece fruto de la gran cantidad de interacciones respecto a la posible irrupción del partido en el parlamento.

El segundo es la poca representación de Unidas Podemos, obteniendo solo una muestra entre las veinte primeras y encontrando a su líder en la posición 74. En este caso, mantiene la misma dinámica que en los debates. Se podría resumir en que interactúa menos y peor con la red política que el resto de líderes de partido.

Tabla 3. Clasificación de actores más importantes por subred en función de los parámetros evaluados durante el día de las elecciones

Subred	Usuarios más importantes	Interacción	Actividad	Centralidad de intermediación	Pagerank
PSOE	PSOE	30	2º (17)	13974	0,011
	sanchezcastejon	51	3º (4)	3º (12671)	3º (0,019)
	miqueliceta	5	11	1265	0,0009
PP	populares	3º (58)	2	1º (124)	1º (0,025)
	pablocasado_	56	1	2º (51)	2º (0,02)
	ppandaluz	5	6	34	0,0008
Ciudadanos	Ciudadanos	1º (71)	1º (17)	10718	0,013
	Albert_Rivera	2º (66)	0	0	0,01
	CsCastillayLeon	3	16	245	0,0007
Unidas Podemos	ahorapodemos	39	4	3169	0,008
	Pablo_Iglesias_	9	1	38	0,001
	pilar_lima	8	2	291	0,003
VOX	vox_es	57	6	3032	0,014
	Santi_ABASCAL	16	2	38	0,0048
	voxnoticias_es	12	8	1029	0,003

Fuente: Elaboración propia

Figura 58. Los tres actores con mayor valoración en los cuatro parámetros durante el día de las elecciones



Fuente: Elaboración propia

A lo largo del 28 de abril se han logrado recoger 4386 tweets de 827 usuarios diferentes.

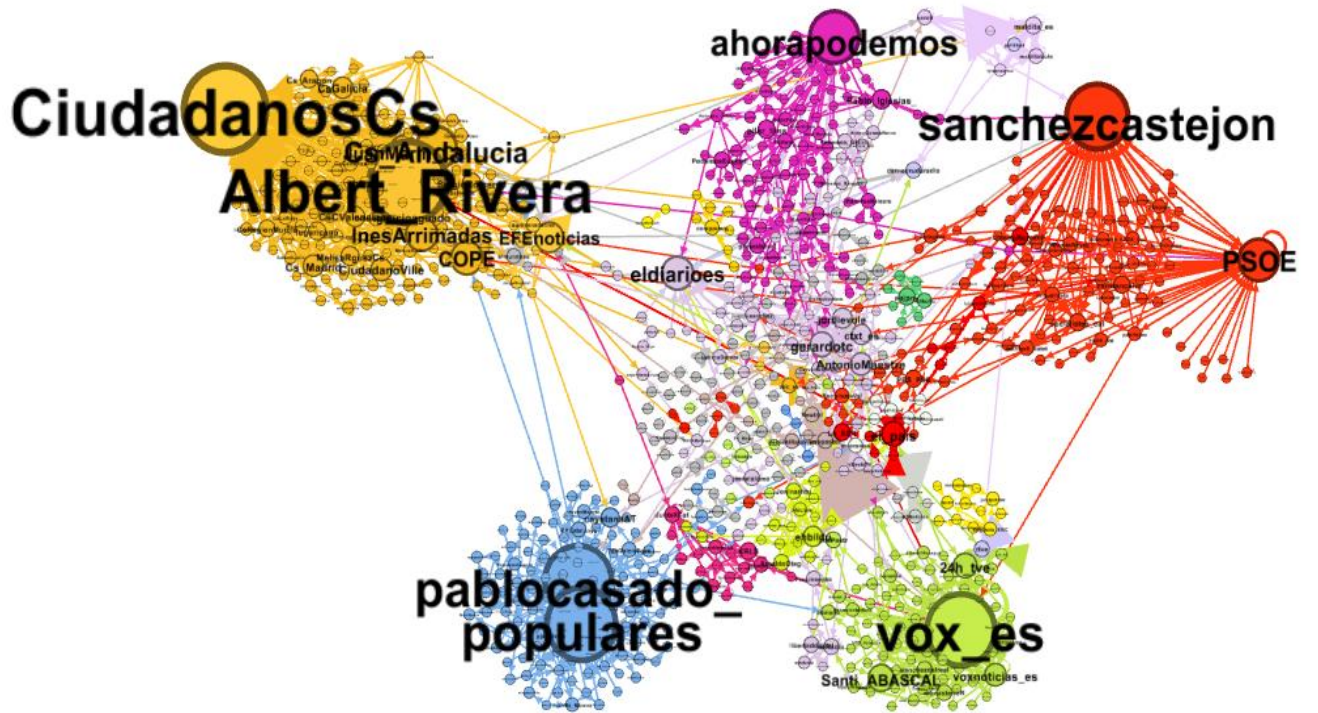
Durante el día de las elecciones, el momento en el que más volumen de tweets se detectó fue en torno a las 00:00, aunque de 22:00 a 23:00 también se detecta un aumento del nivel de interacción. Estos períodos de tiempo coinciden con los primeros resultados oficiales y los resultados finales de las elecciones.

A destacar, las cinco principales cuentas institucionales de partido muestran un alto grado de actividad. PSOE, Unidas Podemos y sobre todo Ciudadanos hacen especial hincapié en la importancia del nivel de participación de cara al resultado electoral. Por otro lado, PP destaca con la publicación en la que se muestra como la diputada por Barcelona, Cayetana Álvarez de Toledo es increpada yendo a votar. Por último, VOX, fiel a su estilo, mantiene ese mensaje cargado de épica augurando un resultado histórico. Parece de los pocos momentos en los que los partidos se olvidan del conflicto catalán para centrarse en el nivel de participación.

Un fiel reflejo de lo que ocurriría finalmente es la contraposición entre Ciudadanos y Unidas Podemos durante el 28 de abril. Ciudadanos acaba logrando un tercer lugar, muy cerca del PP y, aunque sin opciones de gobernar, con la sensación de poder liderar la oposición. Sin embargo, Unidas Podemos obtiene un resultado negativo, similar a la tendencia que había tomado el partido en cuanto a nivel de interacción y prestigio.

Por último, sorprende que, a pesar de mostrar una situación parecida durante los tres eventos, el resultado para PSOE y PP sea dispar, permitiéndole al primero la posibilidad de formar gobierno y ganar las elecciones holgadamente y al segundo obtener el peor resultado de su historia.

Figura 56. Grafo referente al día de las elecciones



Fuente: Elaboración propia

16. Caso de estudio: Polaridad y tendencia de la red política española durante el período de campaña electoral

Como se ha visto en [15], el análisis de sentimiento de texto es un campo interesante a fin de conocer ciertos patrones en determinados usuarios de la red social.

En este caso, la red política española ha sido sometida a una serie de análisis durante determinados eventos. En ellos, los distintos integrantes de la red se han limitado a transmitir información durante el rango de tiempo que abarcaba el suceso en cuestión.

Esta información ha sido difundida en forma de tweets a través de la plataforma y es posible analizar su contenido de distinto modo.

Una forma es a través de las técnicas basadas en *machine learning*, dónde se utilizan modelos de aprendizaje supervisado, es decir, que requiere un conjunto de entrenamiento para conseguir un determinado nivel de predicción y así clasificar la polaridad de textos.

Otra forma, la que se ha utilizado en este trabajo y la que se cita en [15], es la que hace uso de léxicos, es decir, un conjunto de recursos lingüísticos con una semántica definida adecuadamente. De esta forma se consigue determinar la polaridad de un mensaje.

El objetivo de este estudio es lograr entender la tendencia de la red política española a lo largo de la campaña electoral utilizando el cálculo de la polaridad como base. Para ello se ha utilizado el segundo método comentado para la clasificación de textos en función de su polaridad. En este caso, el texto analizado ha sido cada tweet recogido que formase una interacción entre los integrantes de la red. El proceso de cálculo de polaridad ha sido descrito en el apartado 13.

El cálculo se ha realizado por cada evento descrito a lo largo del trabajo. Por cada uno de ellos, se ha calculado el comportamiento de usuarios hacia otros de la red (en función de su polaridad como nodo origen) y la popularidad (grado de polaridad como nodo destino de la interacción). Para la elaboración de este conjunto de datos sólo se han tenido en cuenta a usuarios directamente relacionados con el partido, es decir, cuentas institucionales y de representantes.

16.1. Debate electoral en RTVE el 22 de abril de 2019

El primer evento es el debate de RTVE, del cual se han extraído algunos datos.

Si hablamos de comportamiento, se han tenido en cuenta 176 usuarios con polaridad no neutra, es decir, distinta a cero. Por lo tanto, tan solo el 23,88% de los usuarios ha mostrado un grado de sentimiento en sus interacciones.

De estos 176 usuarios, 73 forman parte del PSOE (41,47%), 61 del PP (34,65%), 10 de Unidas Podemos (5,68%) y 2 de VOX (1,13%). En este caso, Ciudadanos no cuenta con ningún usuario que muestre cierto nivel de polaridad. Del resto, 10 son medios de comunicación (5,68%) y 2 son periodistas (1,13%).

Del total de usuarios, 163 muestran un comportamiento positivo, mientras que 13 son negativos (Tabla 4).

Como se puede apreciar en la Figura 57, destacan positivamente las cuentas del Partido Popular en Melilla, Ceuta y Huelva. Negativamente sobresalen Juan de Dios Ávila, cabeza de lista en Guipúzcoa por VOX, la periodista Cristina Seguí, colaboradora de OkDiario e Intereconomía entre otros y la cadena de radio COPE.

En cambio, si nos centramos en la popularidad de los usuarios en función del grado de polaridad, se cuenta con 54 usuarios, es decir el 7,32% del total. El PSOE cuenta con 20 usuarios (37,03%), 10 el PP (18,51%), 4 Unidas Podemos (7,40%) y 1 VOX (1,85%). Ciudadanos vuelve a quedar sin representación. Entre el total, aparecen 13 medios de comunicación y 1 periodista.

Del total de usuarios, 48 muestran un comportamiento positivo, mientras que 6 son negativos.

Los usuarios más populares son, en primer lugar y catorce puntos por encima del segundo, la cuenta oficial del Partido Popular. En segundo lugar, la cuenta oficial del PSOE y tercero la cuenta del Partido Popular en Burgos.

Los menos populares son el representante en el parlamento europeo de VOX, Hermann Tertsch, Cristina Seguí y la cadena COPE.

Tabla 4. Datos evaluados sobre la polaridad por cada evento

Evento	Tipología polaridad	Nº usuarios	Positivo/a	Negativo/a	Nº Medios
Debate RTVE	Comportamiento	176	163	13	10
	Popularidad	54	48	6	13
Debate Atresmedia	Comportamiento	168	122	46	7
	Popularidad	48	40	8	7
28A – Día de las elecciones	Comportamiento	123	122	1	15
	Popularidad	52	51	1	19

Fuente: Elaboración propia

Figura 57. Clasificación de actores en función de la popularidad y comportamiento



Fuente: elaboración propia

16.2. Debate electoral en Atresmedia el 23 de abril de 2019

El segundo evento es el debate de Atresmedia, dónde se ha repetido el mismo proceso. En este caso, 168 usuarios han mostrado un grado de comportamiento, es decir, un 21,9% del total. De estos, 66 pertenecen a la red de Ciudadanos, 36 a la del PSOE, 32 a la del PP y 11 a la de Unidas Podemos. Sorprenden dos aspectos, el primero, la irrupción de Ciudadanos y segundo, la desaparición de VOX. En esta lista participan 7 medios de comunicación y 2 periodistas.

Del total, 122 son considerados usuarios positivos y 46 negativos.

Con mejor comportamiento destacan la cuenta del Partido Popular en Melilla y dos diputadas del Partido Popular, Ana Madrazo y Carmen Riobos.

Los usuarios con peor comportamiento son la ministra de economía y empresa, Nadia Calviño, y las cuentas de Ciudadanos en Santa Cruz de Tenerife y Castilla y León.

Por otro lado, se dispone de 48 usuarios en relación a la popularidad encontrada en las interacciones. Esto supone un 6,25% del total de usuarios que interaccionaron durante el debate. En este caso, 10 forman parte de la red del PSOE, 8 de la de Ciudadanos, 7 del PP, 5 de Unidas Podemos y 1 de VOX. Al igual que en el análisis de comportamiento, aparecen 7 medios de comunicación y 2 periodistas.

De estos, 40 son positivos y 8 negativos. Los más populares son, 20 puntos por delante del segundo, la cuenta oficial del Partido Popular. En segundo lugar, la cuenta oficial de Ciudadanos y en tercero la cuenta oficial de las nuevas generaciones del Partido Popular en España.

Los tres menos populares son, muy destacado del resto, Albert Rivera, la ministra de justicia, Lola Delgado y el secretario general de Unidas Podemos en Castilla y León, Pablo Fernández.

16.3. 28 de abril – Día de las elecciones

Por último, en el día de las elecciones, aparecen 123 usuarios con cierto grado de comportamiento (14,87%). Del total, 40 son parte de la red del PSOE, 15 de la de VOX, 14 de Ciudadanos, 13 del PP y 3 de Unidas Podemos. También forman parte 15 medios de comunicación y 5 periodistas.

De estos usuarios, 122 mantienen un comportamiento positivo y tan solo 1 de forma negativa.

Los actores con mejor comportamiento son la periodista Ana Pastor, y los medios de comunicación Diario Vasco y Antena3. Este último, ligeramente por delante de otro medio como el periódico ABC.

En este caso, debido a la poca representación de usuarios con comportamiento negativo, solo se puede citar la cuenta oficial de EHBildu.

Por otra parte, aparecen 52 usuarios con grado de popularidad distinto a cero. Esto significa un 6,28%. Del total, 12 pertenecen a la red del PSOE, 4 a la del PP, 3 son de VOX y Ciudadanos y Unidas Podemos empatan con 2. Se encuentran 19 medios de comunicación y 2 periodistas.

Se consideran positivos 51 usuarios y, como en el grado de comportamiento, solo 1 aparece con valor n negativo.

Los más populares son la cuenta oficial del Partido Popular, Newtral y el Diario Vasco. Al contrario, se vuelve a encontrar la cuenta de EHBildu como el único usuario con popularidad negativa.

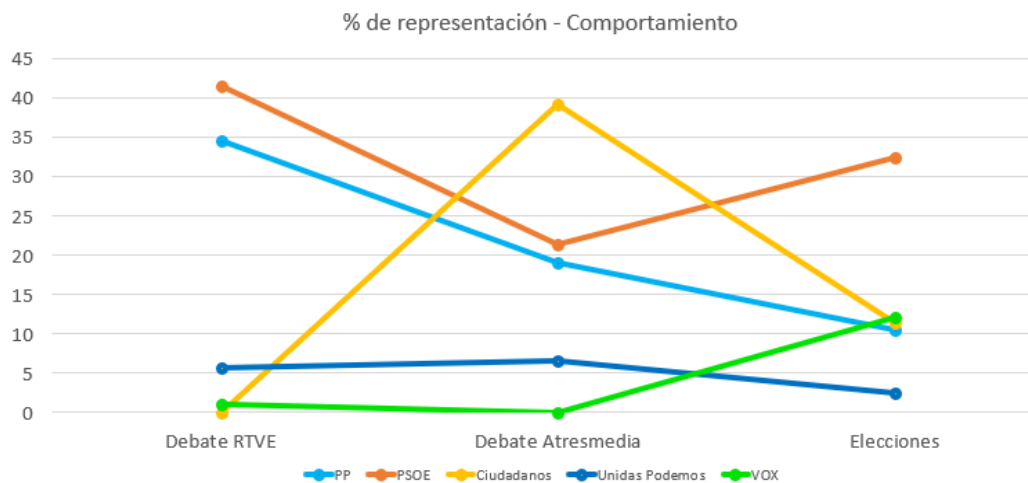
16.4. Tendencia de la red política española

Se pueden afirmar dos cosas. La primera es que la representación, tanto de comportamiento como de popularidad, siguen una misma tendencia. Ningún partido experimenta ningún cambio significativo de comportamiento a popularidad.

La segunda es la sorprendente similitud de la tendencia de polaridad de Ciudadanos respecto a su grado de interacción a lo largo de la campaña. Se recuerda que Ciudadanos no muestra apenas representación en el debate de RTVE, la única cuenta que llamaba la atención debido a su alto PageRank fue la oficial del partido. Sin embargo, en los demás eventos muestra otra imagen más activa y dominadora.

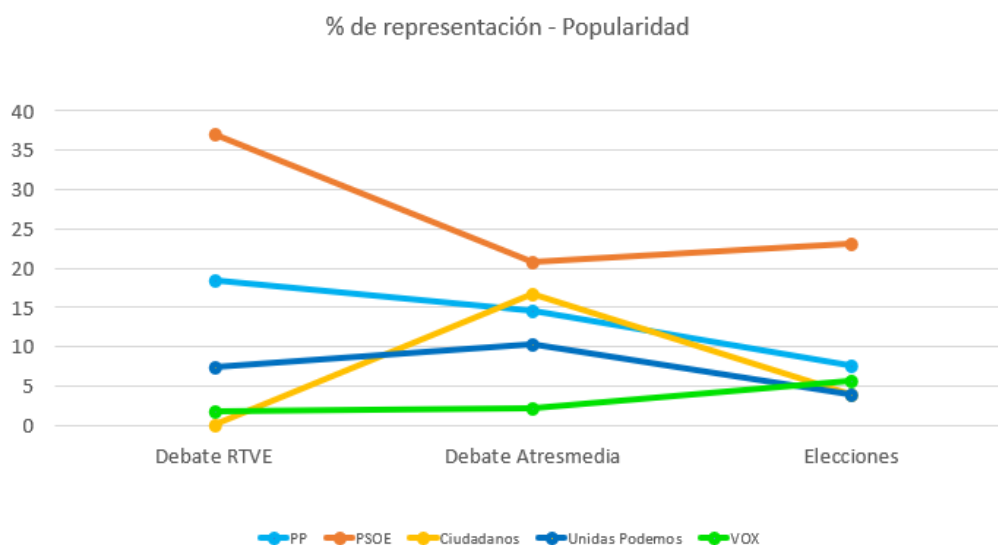
Se puede observar visiblemente lo expuesto en la Figuras 58 y 59, dónde se aprecia cómo evoluciona a lo largo de los tres eventos el porcentaje de representación de cada partido sobre el total de usuarios que muestran cierto comportamiento y popularidad.

Figura 58. Evolución del porcentaje de representación de comportamiento por partido



Fuente: elaboración propia

Figura 59. Evolución del porcentaje de representación de popularidad por partido



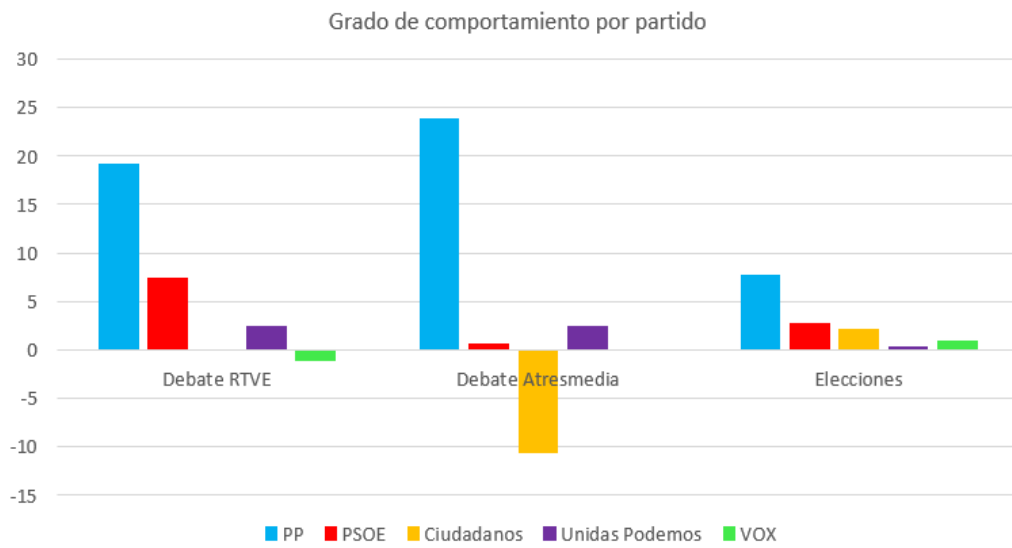
Fuente: elaboración propia

Si se fija la atención en las figuras 60 y 61, es palpable el dominio del Partido Popular sobre el resto de partidos referente al grado de polaridad obtenido. Es el único partido que se mantiene en cabeza durante los tres eventos, además de su abrumadora diferencia respecto al resto de fuerzas.

Destaca la posición de Ciudadanos que, a pesar de aportar gran cantidad de interacciones a partir del segundo evento e inclusive, interacciones con cierto grado de polaridad, tanto el grado de comportamiento como el de popularidad del partido acaban siendo negativos en el debate de Atresmedia. Como se puede observar en las Figuras 58 y 59, el día de las elecciones, vuelve detentar un menor porcentaje de representación y, llamativamente, su grado de polaridad vuelve a ser positivo. Esto indica que cuanto más interacciona la comunidad de Ciudadanos, peor consideración tiene para el resto de usuarios de la red.

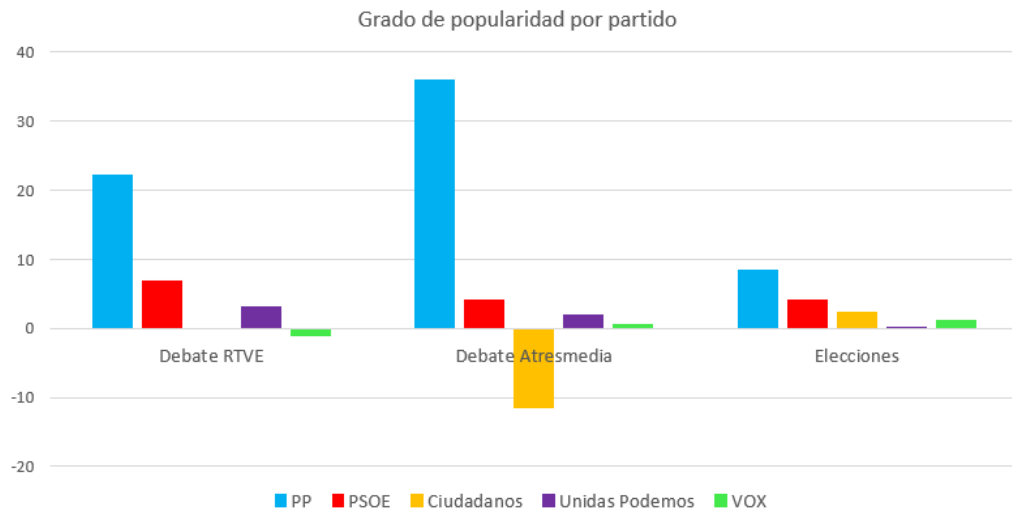
Por último, también es interesante la analogía que muestra la tendencia de Unidas Podemos y VOX respecto al posicionamiento y prestigio en la red durante la campaña. Recapitulando, a partir del segundo debate, la comunidad de Unidas Podemos comienza a perder importancia en la red a la par que la de VOX adquiere cierto prestigio. La segundo es razonable puesto que VOX no participó en los debates, pero no tanto la tendencia que sigue Unidas Podemos. El descenso en el grado de interacción, el prestigio en la red o la polaridad del contenido de su mensaje son solo algunos ejemplos de la decadencia del partido de Pablo Iglesias.

Figura 60. Evolución del grado de representación de comportamiento por partido



Fuente: elaboración propia

Figura 61. Evolución del grado de representación de popularidad por partido



Fuente: elaboración propia

Si se tiene en cuenta la inclusión de los medios de comunicación en las comunidades de los partidos, aparecen nuevos resultados.

En el debate de RTVE, de los diez medios de comunicación que ofrecieron un grado de comportamiento, dos de ellos se posicionan en la red de Unidas Podemos (ABC y El Confidencial) y otros dos en la de PSOE (El País) y VOX (El Mundo).

Si hablamos de popularidad, de los trece medios, tres se posicionan en la red de Unidas Podemos (ABC, El Diario y El Confidencial), uno al PSOE (El País) y otro a VOX (El Mundo).

En el debate de Atresmedia, de los siete medios, uno forma parte de la red del PP (Antena 3), uno a la del PSOE (El País) y otro a la de Unidas Podemos (ABC).

En cuanto a la popularidad, de los siete medios de comunicación que ofrecieron un grado de comportamiento, dos de ellos se posicionan en la red del PSOE (El País y El Diario), dos en la del PP (La Vanguardia y Antena 3) y otra en la de Unidas Podemos (ABC).

Por último, el 28 de abril, de los quince medios de comunicación que ofrecieron un grado de comportamiento, cuatro de ellos se posicionan en la red de Ciudadanos (ABC, 20 Minutos, El Español y la cadena COPE), uno en la de Unidas Podemos (Periódico Castilla-La Mancha) y otro en la de VOX (OkDiario).

Haciendo referencia a la popularidad, de los diecinueve medios, tres se posicionan en la red de Ciudadanos (ABC, 20 Minutos, El Español) y uno en la del PSOE (El Huffington Post).

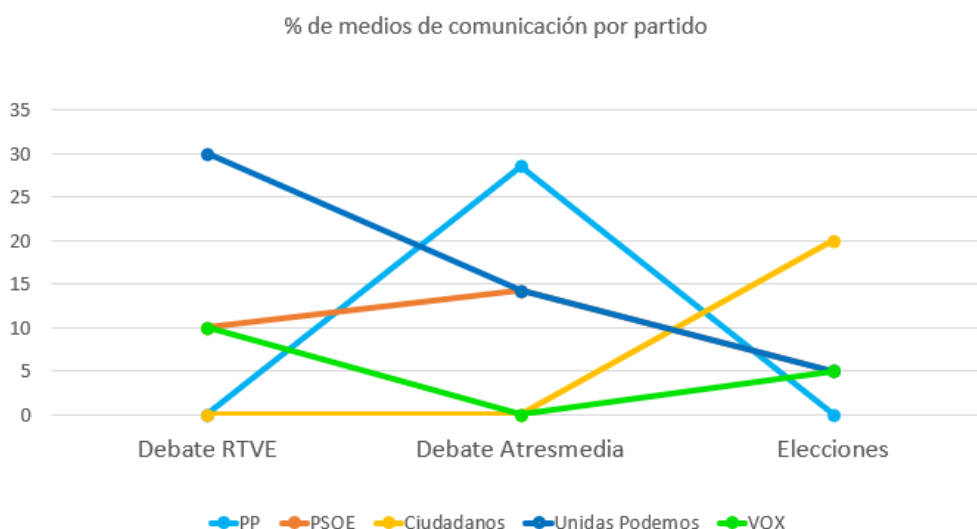
Observando el contenido de la interacción entre estas cuentas no observamos un alto grado de sesgo en su contenido. Aunque existen excepciones, por ejemplo con VOX y OkDiario en la que el partido de Santiago Abascal comparte un tweet del medio digital el 28 de abril en el que se hace especial atención a la desigualdad en cuanto a la relación de voto-escaño entre VOX y los partidos independentistas.

Por lo tanto, se puede afirmar que el posicionamiento de los medios de comunicación en la red, no es principalmente por afinidad ideológica sino por el nivel de interacción con los integrantes en esta.

Obviando la posición del usuario mientras interactúa, es decir, si cómo nodo origen o destino, se pueden exponer algunos apuntes. Como indica la Figura 62, lo primero que se observa es la poca representación de medios de comunicación que posee VOX. Esto se refleja en su consecuente valor bajo de polaridad aportado por los medios afines a este partido, como se aprecia en la Figura 63.

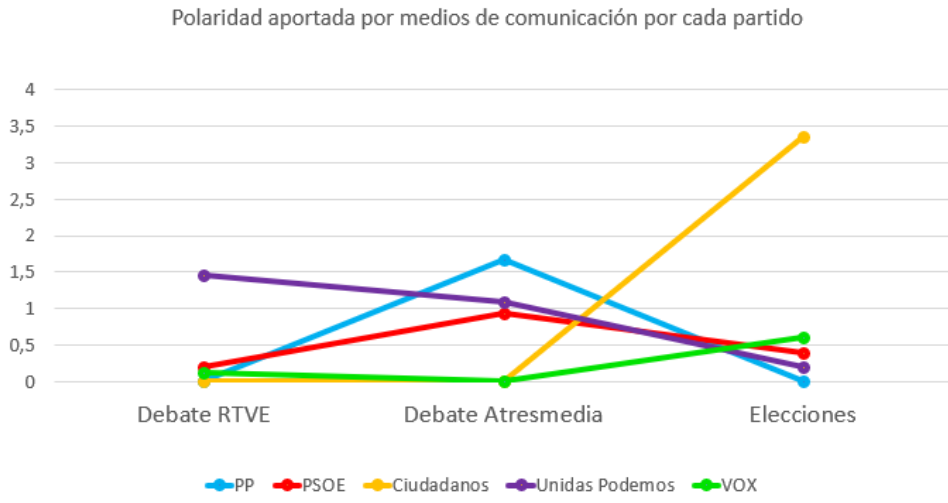
También se manifiesta que el valor aportado por los medios de comunicación sigue la misma tendencia que el resto de usuarios, aunque algunos partidos resulten más beneficiados que el resto. Por ejemplo, al contrario que VOX, Ciudadanos acaba logrando una puntuación tres puntos por encima del resto, lo que supone un 600% más que el segundo. Aunque la puntuación sea baja, no deja de ser la representación de un solo evento. Si ponemos de relieve estos datos y ampliamos la perspectiva, el poder que aplican los medios de comunicación sobre cualquier usuario o grupo de usuarios, puede llegar a ser un aspecto clave en las campañas de los distintos partidos políticos. De hecho, no es inusual que cierto partido político salga perjudicado por la interacción de un medio. Un ejemplo de ello es Antena 3 durante el debate de RTVE. En dicho evento, Antena 3 se posiciona en la comunidad del PP. Sumando todos los grados de popularidad en base a la polaridad como nodo destino, la cadena de televisión logra una valoración de -0,275, algo que, en cierto modo, podría tener relevancia en la evolución del partido. Si se posiciona en la red del PP significa que existe algún tipo de relación. Dependiendo de la naturaleza de la relación esta polaridad puede influir en el partido o no. En el caso de que sea meramente informativa y se haya posicionado así debido al contenido del mensaje, no tendría por qué afectar. Por otro lado, si existe información sesgada en sus tweets orientada hacia una ideología de derechas, el medio de comunicación podría ser vinculado con el partido por los usuarios. Por lo tanto, si RTVE adquiere una popularidad negativa, sería algo negativo para el Partido Popular, en este caso.

Figura 62. Porcentaje de representación de medios de comunicación por partido



Fuente: elaboración propia

Figura 63. Polaridad aportada por los medios de comunicación por partido



Fuente: elaboración propia

17. Discusiones

A lo largo del trabajo se han expuesto una serie de datos obtenidos a partir de la metodología empleada. A partir de ella se pueden contextualizar los resultados alcanzados.

Durante el desarrollo del trabajo se ha hecho hincapié en cinco partidos políticos dada la relevancia de estos. En cualquier caso, todo usuario integrante de la red ha sido útil para la consecución de los objetivos.

El Partido Popular ha mantenido una actuación lineal, es decir, se ha mantenido, salvo excepciones, en una posición privilegiada en todos los aspectos estudiados de principio a fin del proyecto. Es la fuerza que más aliados no políticos posee. Como se ha visto a través del estudio de 'follows' de cada usuario, se puede afirmar que la mayoría de la prensa escrita se posiciona en su comunidad. No se puede obviar que las fuerzas y organismos públicos, en su mayoría, muestran la misma conducta. Como excepción, durante el 28 de abril, se percibe una notoria falta de comunicación con el resto de partidos. En este evento, el partido no necesita un usuario posicionado estratégicamente con un grado de centralidad alto debido a que prácticamente no existen interacciones con el resto de usuarios no pertenecientes al partido.

Por su parte el PSOE, muestra una postura parecida pero siempre por debajo del Partido Popular. Sin duda, si se tuviese que elaborar una clasificación respecto a la totalidad de métricas establecidas, quedaría en segundo lugar, alejado tanto del Partido Popular como del resto de partidos. Quizás, junto con Ciudadanos sea el partido que posee mayor fidelidad de medios de comunicación. Tanto El País como La SER son habituales en la comunidad del partido socialista. Por otro lado, aunque en función de los 'follows', es el único partido capaz de posicionarse en una de las 5 grandes comunidades a Coalición Canaria, que durante los tres eventos no interactuó con ningún otro usuario externo a su comunidad. Recordemos que durante el análisis estático se posicionó en la red del PSOE a Coalición Canaria, conocida por su ambigua ideología política, ya que no suele posicionarse claramente hacia ninguna vertiente. Esto adquiere más valor e interés si observamos que durante los tres eventos analizados durante el proyecto, Coalición Canaria no muestra comunicación alguna con algún integrante fuera de su partido.

En el caso de Ciudadanos, es el partido con mayor cohesión entre sus integrantes. Esto puede ser debido, en gran medida, al tamaño de la red. Cuanto más pequeña sea una red, más posibilidades existen de que la relación entre sus integrantes sea más fuerte y por lo tanto la cohesión de la red sea mayor. En el grafo estático, Ciudadano es el quinto partido en número de integrantes con un 12,61% de la red, muy cercano a VOX (13,37%), pero muy lejano del resto: PSOE (18,45%), PP (22,18%) y Unidas Podemos (30,65%). Al igual que el Partido Popular, los miembros de la comunidad del partido de Albert Rivera no necesitan uno o varios usuarios como intermediarios para establecer una relación fuera de su comunidad. Sin embargo, no por el mismo motivo, sino porque los propios integrantes del partido suelen ser

autosuficientes, es decir, interactúan directamente fuera del partido. Es destacable la mala popularidad de Albert Rivera durante el debate de Atresmedia y, por consiguiente, del partido. Se llega a la conclusión de que cuanto mayor es el grado de interacción del partido, mayor porcentaje de polaridad negativa existe en sus interacciones. Si hacemos uso de nuestro laboratorio de datos y observamos algunos tweets con polaridad negativa en los que actúan como nodo de destino (popularidad) Albert Rivera o la cuenta oficial de Ciudadanos aparecen como nodo origen integrantes de las comunidades PSOE y VOX, salvo las interacciones provenientes de La Sexta, que forma una comunidad propia durante el debate de Atresmedia. Esto es entendible dada la situación política de ambos partidos. Ciudadanos tiene como objetivo llamar la atención de aquellos votantes indecisos afines al PSOE, por lo que la comunidad del partido socialista lleva a cabo la misma estrategia por el mismo motivo. Pero también existe otra causa, aunque menos común, que aumenta este grado de polaridad negativa y es la retroalimentación de los integrantes de Ciudadanos de su propio contenido. Como se ha descrito durante el trabajo, Ciudadanos posee la red más cohesionada y por lo tanto, parece más común que interactúen entre ellos. Si el contenido de su mensaje adquiere un grado de polaridad negativo y se generan interacciones con este, hace que se genere un mayor grado de popularidad negativa en el caso de Albert Rivera o la cuenta de Ciudadanos.

De Unidas Podemos se pueden establecer varias conclusiones. La primera es la tendencia negativa del partido durante el proceso electoral. Es visible el empeoramiento en cuanto a centralidad y prestigio de la red en su totalidad. La segunda es que no es el partido que menos interacciona con el resto de usuarios de la red política, pero suele obtener un valor muy bajo de prestigio. Por lo tanto, parece evidente que interactúa con actores de su propia red o no muy relevantes fuera de ella, manteniendo un cierto grado de aislamiento respecto del resto de partidos fruto, seguramente, de una ideología cargada de propuestas que solo Unidas Podemos pone encima de la mesa, distinguiéndose así del resto de fuerzas.

Respecto de VOX, es el antagonista de Unidas Podemos, durante el transcurso de los debates permanece un tanto al margen del resto de partidos, pero el día de las elecciones muestra una imagen más realista. Se mantiene en una quinta posición, pero lejos de la sensación de partido minoritario que mostraba en los dos primeros eventos. También es destacable que es el partido que menos medios promedia en su red durante los tres eventos, pero es capaz de acoger a grandes nombres como El Mundo durante el 28 de abril.

Sobre el resto de partidos, las fuerzas independentistas visualmente en la red un posicionamiento siempre cerca de la izquierda española, sobre todo de Unidas Podemos. Aunque en ocasiones, como en el caso del debate de RTVE, exceptuando EAJPNV, el resto se posicionan entre Unidas Podemos y Ciudadanos. Tiene sentido si se recuerda que Ciudadanos nace como partido autonómico en Cataluña y tiene fuertes vínculos en contra del independentismo catalán. Por lo tanto, esta cercanía, lejos de ser una muestra de afinidad, representa la contrariedad ideológica entre ambas fuerzas.

El resto de partido, es decir, EAJPNV, EnMarea, Compromís y Coalición Canaria, no muestran interacción alguna con el resto de la red. Se limitan a interactuar dentro de su propia comunidad.

Por último, se ha visto como los medios de comunicación no suelen tender claramente hacia ningún partido. Existen excepciones como las del PSOE con La SER o Ciudadanos con El Español, pero la mayoría de ellos se mantienen al margen, creando su propia comunidad. De todos modos, es evidente que existen afinidades entre medios y partidos y, como se ha visto con el estudio de la polaridad, pueden llegar a ser un aspecto clave durante la campaña electoral.

18. Conclusiones

Una vez concluido el trabajo, es hora de volver la mirada y hacer autocrítica sobre lo conseguido.

Se ha elaborado la red política española a partir de unos criterios expuestos al comienzo del proyecto. Se ha conseguido realizar una serie de pruebas sobre la red logrando así, información útil para la detección de los usuarios y sucesos más relevantes durante la campaña electoral. Por último, se ha propuesto un caso de estudio en el que se ha observado como puede llegar a influir el contenido compartido en Twitter por los integrantes de la red en su objetivo político, así como la influencia de los medios de comunicación en cada partido.

Por lo tanto, se puede afirmar que, en general, los objetivos propuestos al comienzo del trabajo han sido cumplidos satisfactoriamente, aunque siempre mejorables.

Existen varios puntos de mejora para la consecución del trabajo. El primero es la experiencia; es evidente que cuanto más tiempo se empeña en desarrollar un trabajo sobre una materia, más se conoce de ella y mejores resultados se obtienen. Esto implica tanto conocimiento teórico sobre análisis de redes sociales como la destreza en la práctica, en la que se incluyen conocimientos sobre programación.

El segundo es el alcance de los recursos utilizados para la consecución del trabajo. Con un equipo más potente, la extracción y almacenamiento de datos en Twitter habría sido mucho más sencilla.

El tercero es la cantidad de datos recogidos. Con un mayor volumen de tweets, se podrían haber obtenido resultados más exactos, sobre todo en materia de análisis de sentimiento. Esto iría ligado también al alcance de los equipos utilizados.

Por último, este trabajo abarca el período de campaña electoral y los días posteriores al día de las elecciones. Si el trabajo se hubiese desempeñado abarcando un período de tiempo más amplio, se podrían haber obtenido datos que revelasen de forma más exacta el efecto del uso de redes sociales en cada partido político en su objetivo electoral, así como la detección más precisa de movimientos relevantes por parte de los integrantes de la red durante el período establecido. Por ejemplo, si se ampliase el período hasta el proceso de investidura, si es que se llega a producir en primera instancia. En caso contrario, serviría para contrastar la red en dos campañas electorales de la misma naturaleza.

A partir de los objetivos alcanzados en el trabajo, se plantean futuras líneas de investigación como, por ejemplo:

- Evolución del prestigio en la red de las cinco grandes fuerzas a partir del período de investidura.
- Consecución de los mismos objetivos para las elecciones municipales y europeas.

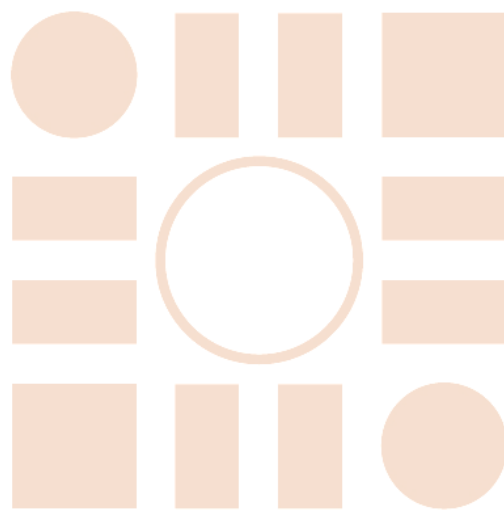
- Contraste de los resultados obtenidos en elecciones generales respecto a las municipales y europeas.
- Construcción de algoritmo para la detección de análisis de sentimiento de cada partido político y la detección de los usuarios más influyentes en la red.

19. Bibliografía

- [1] J. A. Barnes, "Class and committees in a Norwegian island parish", Human Relations, vol. 1, no. 7, pp. 39-58, 1954.
- [2] PJ Carrington, J Scott, S Wasserman, Models and methods in social network analysis. United States: Cambridge University Press, 2005.
- [3] O. Ore, Grafos y sus aplicaciones. Madrid: DLS-EULER, 1995.
- [4] N. L. Biggs, Algebraic Graph Theory, Cambridge: Cambridge University Press, 1974.
- [5] E. Bujalance y otros, Elementos de Matemática Discreta, Editorial Sanz y Torres: 1993.
- [6] C. Madariaga, J. Palacio, PSICOLOGÍA SOCIAL APLICADA Y ANÁLISIS DE REDES SOCIALES. Psicología Social. Teoría y práctica, Ediciones Uninorte. P. 103. En: Angarita, C. (Compiladora), 2007.
- [7] L. Freeman, The development of social network analysis. Vancouver, Canadá: ΣP Empirical Press, 2004.
- [8] Andre, R. Pinheiro, Social Network Analysis in Telecommunications. Hoboken, NJ, USA: John Wiley & Sons, 2011.
- [9] Clemente, Filipe Manuel, F. M. Lourenço Martins, and R. S. Mendes, Social Network Analysis Applied to Team Sports Analysis. 1st Ed. 2016 ed. Cham: Springer International, 2016.

- [10] Tayebi, Mohammad A., y Uwe. Glässer, *Social Network Analysis in Predictive Policing: Concepts, Models and Methods*. United States: Springer, 2016.
- [11] L. Freeman, *Encyclopedia of Psychology*.
- [12] Young, T. V., Wang, Y. & Lewis, W. D, *Explaining How Political Actors Gain Strategic Positions: Predictors of Centrality in State Reading Policy Issue Networks*. Educational Policy. North Carolina State University: Sagepub, 2016
- [13] J. Ebermann, K. Stanoevska-Slabeva, T. Plotkowiak, *A Longitudinal Social Network Analysis of German Politicians Twitter Accounts*. Riva del Garda: Sunbelt XXX Conference, 2010.
- [14] García Ortega C. y Zugasti Azagra, R., *La campaña virtual en twitter: análisis de las cuentas de Rajoy y de Rubalcaba en las elecciones generales de 2011*. Vol. 19. Núm. Especial Febrero. Págs. 299-311. 2014.
- [15] Bruno Á. Souza, Thais G. Almeida, Alice A. F. Menezes, Fabíola G. Nakamura, Carlos M. S. Figueiredo, Eduardo F. Nakamura, *For or Against? Polarity Analysis in Tweets about Impeachment Process of Brazil President*. Teresina, Estado de Piauí, Brasil: ACM. 2016.

Universidad de Alcalá
Escuela Politécnica Superior



ESCUELA POLITECNICA
SUPERIOR



Universidad
de Alcalá