

Trabajo fin de grado

Formación de opiniones en las redes sociales e influencia de los medios de comunicación.



Manuel Suárez Román

Escuela Politécnica Superior
Universidad Autónoma de Madrid
C\Francisco Tomás y Valiente nº 11

UNIVERSIDAD AUTÓNOMA DE MADRID
ESCUELA POLITÉCNICA SUPERIOR



Grado en Ingeniería Informática

TRABAJO FIN DE GRADO

**Formación de opiniones en las redes sociales e
influencia de los medios de comunicación.**

Autor: Manuel Suárez Román

Tutor: Simone Santini

abril 2023

Todos los derechos reservados.

Queda prohibida, salvo excepción prevista en la Ley, cualquier forma de reproducción, distribución comunicación pública y transformación de esta obra sin contar con la autorización de los titulares de la propiedad intelectual.

La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (*arts. 270 y sgts. del Código Penal*).

DERECHOS RESERVADOS

© 19 de abril de 2023 por UNIVERSIDAD AUTÓNOMA DE MADRID
Francisco Tomás y Valiente, nº 1
Madrid, 28049
Spain

Manuel Suárez Román

Formación de opiniones en las redes sociales e influencia de los medios de comunicación.

Manuel Suárez Román

C\ Francisco Tomás y Valiente N° 11

IMPRESO EN ESPAÑA – PRINTED IN SPAIN

Quizá haya enemigos de mis opiniones, pero yo mismo, si espero un rato, puedo ser también enemigo de mis opiniones.

Jorge Luis Borges

AGRADECIMIENTOS

Quiero agradecer a mi tutor, Simone Santini, el trabajo de orientación, sobresaltos incluidos, que ha realizado a lo largo de estos meses.

A mis amigos, a los que habéis estado a mi lado y habéis formado parte de mi etapa universitaria.

A mi padrino, Mariano, por haber despertado siempre ese espíritu de inquietud en mí.

Pero sobre todo, a mi familia, en especial a mis padres, Justina y Miguel, por haberme apoyado en todas las decisiones que he tomado a lo largo de estos años.

De todo corazón, gracias.

RESUMEN

En 2018 la compañía inglesa Cambridge Analytica saltó a la fama por sus vinculaciones con la campaña presidencial de Donald Trump y con el Brexit. Este hecho colocó en un primer plano la importancia de los datos de los usuarios de las redes sociales, y puso de manifiesto lo fácilmente influenciable que puede resultar la opinión de los mismos. Pese a todo esto, los diversos estudios realizados en el campo de la formación de opiniones no incluyen, o lo hacen únicamente de una manera muy superficial, los condicionantes que las redes sociales aportan a las interrelaciones humanas.

En este trabajo en primer lugar adaptamos a una estructura de grafo (topología característica de las redes sociales), uno de los principales modelos desarrollados de formación de opiniones. Después, estudiamos los diferentes condicionantes que hacen que la opinión de los componentes de dicha red transite desde el consenso hasta la polarización en posiciones extremas. Y vemos que el principal factor determinante para esta variación es el apego que cada individuo tiene a su propia opinión. A este apoyo a lo largo del desarrollo de esta memoria denominaremos como λ . También comprobamos que resulta relevante la influencia que cada uno de los agentes de la red pueda tener sobre cada individuo en un determinado momento.

Para completar el trabajo desarrollado, estudiamos la incorporación de una serie de nodos a nuestro grafo que representan a los medios de comunicación en las redes sociales. Con este fin, proponemos un algoritmo que determina con qué medios de comunicación se conectará cada individuo de la red en función de la opinión de estos y la propia. Comprobamos así que el papel de los medios de comunicación en la formación de las opiniones es el de suavizar las situaciones, haciendo que no se llegue ni al consenso absoluto ni a posiciones de alta polarización.

PALABRAS CLAVE

Redes sociales, grafos, formación de opinión, medios de comunicación

ABSTRACT

In 2018, the British company Cambridge Analytica became famous for its links to Donald Trump's presidential campaign and Brexit. This brought to the forefront the importance of social media users' data, and highlighted how easily their opinions can be influenced. Despite all this, the various studies carried out in the field of opinion formation do not include, or do so only superficially, the conditioning factors that social networks bring to human interrelations.

In this paper we first adapt one of the main developed models of opinion formation to a graph structure (topology characteristic of social networks). Then, we study the different determinants that make the opinion of the components of such a network move from consensus to polarisation in extreme positions. And we see that the main determinant of this variation is the attachment that each individual has to his or her own opinion. Throughout the development of this report we will refer to this attachment as λ . We also found that the influence that each of the agents in the network may have on each individual at any given time is relevant.

To complete the work developed, we study the incorporation of a series of nodes to our graph that represent the media in the social networks. To this end, we propose an algorithm that determines which media each individual in the network will connect with according to their opinion and his own. We thus prove that the role of the media in the formation of opinions is to soften situations, ensuring that neither absolute consensus nor highly polarised positions are reached.

KEYWORDS

Social networks, graphs, opinion dynamics, media

ÍNDICE

1 Introducción	1
1.1 Motivación	1
1.2 Objetivos y estructura del trabajo	2
2 Estado del arte	3
2.1 Modelos de formación de opiniones	3
2.1.1 El modelo cinético de gases ideales	3
2.1.2 El modelo del votante	5
2.1.3 Formación de opiniones con tolerancia limitada	7
2.1.4 Otros modelos de formación de opiniones	9
2.2 Las redes sociales como grafos	10
2.2.1 El modelo de Leskovec	10
2.2.2 El modelo de Erdős-Rényi	11
2.2.3 El modelo de Barabási-Albert	12
2.3 La función de distribución de probabilidad beta	13
3 El modelo cinético de gases ideales y las redes sociales	15
3.1 El modelo de Leskovec y la evolución de la opinión media	16
3.2 Estudio de la evolución de la opinión en función del parámetro λ	20
3.3 Otros modelos de creación de grafos	22
3.4 Otras consideraciones	24
4 El papel de los medios de comunicación en la formación de opiniones	29
4.1 Diseño de algoritmo para conexión a medios de comunicación	29
4.2 Precisión del modelo	32
4.3 Valor de los parámetros de la función beta	34
4.4 Resultados obtenidos	35
5 Conclusiones y limitaciones	39
5.1 Conclusiones	39
5.2 Limitaciones y trabajo futuro	40
Bibliografía	42
Apéndices	43
A Resultados de la encuesta realizada	45

A.1	El estudio de Ad Fontes Media	45
A.2	Posicionamiento de los medios de comunicación	46
A.3	Resultados de la encuesta	46

LISTAS

Listado de algoritmos

3.1	Algoritmo que define la interacción entre nodos condicionados por la tolerancia	18
4.1	Algoritmo para la conexión a medios de comunicación	31

Listado de ecuaciones

2.1	Ecuación para la formación de opiniones propuesta por M. Lallouache	4
2.2	Cálculo de la opinión media	4
2.3	Ratio de cambio de estado de cada votante	6
2.4	Cálculo del hamiltoniano propuesto en el modelo de Ising	6
2.5	Ecuación para la formación de opiniones propuesta por Parongama Sen	9
2.6	Probabilidad de conexión de un nuevo nodo según el modelo de Barabási	12
2.7	Función de densidad descrita por la distribución beta	13
2.8	Función beta	13
2.9	Cálculo de la esperanza de la distribución beta	13
2.10	Cálculo de la varianza de la distribución beta	13
2.11	Cálculo de la esperanza de la distribución beta con mismo valor de parámetros	14
2.12	Cálculo de la varianza de la distribución beta con mismo valor de parámetros	14
3.1	Cálculo del valor del grado medio de los nodos de la red	22
4.1	Cálculo de la DCM para el modelo propuesto de conexión a medios de comunicación	32

Listado de figuras

2.1	Resultados experimentales observados por Lallouache	5
2.2	Resultados experimentales observados por Hegelsmann con valores simétricos de ϵ ..	8
2.3	Resultados experimentales observados por Hegelsmann con valores asimétricos de ϵ ..	8
2.4	Variación de los intervalos de confianza de Hegelsmann	9
2.5	Grafos de Leskovec	11
2.6	Funciones de densidad de probabilidad beta con distintos parámetros α y β	14

3.1	Comportamiento de la opinión media para distintos valores de λ	17
3.2	Distribución de la opinión para distintos valores de λ	17
3.3	Distribución de la opinión en función del valor de la tolerancia	19
3.4	Evolución de la opinión media para distintos valores de tolerancia	20
3.5	Comportamiento de la opinión al aplicar la estructura de Leskovec	21
3.6	Evolución de la opinión media y de la varianza de esta respecto de λ	21
3.7	Comportamiento de la opinión al aplicar la estructura de Erdös-Rényi	22
3.8	Comportamiento de la opinión al aplicar la estructura de Barabási-Albert	23
3.9	Comportamiento de la opinión al aplicar la estructura de un grafo completamente conexo	24
3.10	Valor de λ_c a medida que varía la tolerancia y la conectividad media del grafo	25
3.11	Variación de λ_c en función del número de nodos	25
3.12	Variación de λ_c en función del valor mínimo de ϵ	26
3.13	Comportamiento de la opinión con la estructura de Leskovec y valor mínimo de $\epsilon = -1$	27
3.14	Comportamiento de la varianza de la opinión con valor mínimo de $\epsilon = -1$	28
4.1	Distribución beta estirada y centrada en una opinión	30
4.2	<i>DCM</i> cuando cada individuo se conecta a 4 medios de comunicación	33
4.3	Distancia cuadrática media con diferentes valores de medios de comunicación	33
4.4	<i>DCM</i> del algoritmo propuesto en función del valor de γ	35
4.5	Comportamiento de la opinión al incorporar los medios de comunicación al modelo	36
4.6	Evolución de la varianza de la opinión al incorporar los medios de comunicación	36
4.7	Mediciones sobre el modelo propuesto con el valor mínimo de $\epsilon = -1$	37
A.1	Distribución de los medios de comunicación españoles.	46
A.2	Resultados obtenidos en la primera pregunta de la encuesta realizada.	47
A.3	Resultados obtenidos en la encuesta realizada.	48

Lista de tablas

4.1	Tabla de distancias del modelo presentado en función de los medios de comunicación	34
A.1	Distribución del posicionamiento de los medios de comunicación españoles	47

INTRODUCCIÓN

1.1. Motivación

El mundo actual se caracteriza por disponer de una gran cantidad de información al alcance de la mano de cualquier individuo. En los últimos años se ha experimentado un gran desarrollo de las redes sociales y de la transmisión de la información en las misma. Esto ha llevado, entre otras, a la popularización de las conocidas como *fake news*, que aprovechan la tendencia del ser humano a no contrastar fuentes [10]. La influencia de estas en la población, junto con la que ejercen los medios de comunicación, llega a provocar que grandes masas de individuos cambien su opinión de forma repentina basándose en las interacciones que se producen entre estos.

Pese a la importancia de este asunto, y debido a lo reciente del mismo, no existen demasiados estudios que traten el asunto de la difusión de la información en las redes sociales. Sí que podemos encontrar diversos ensayos que versan sobre la evolución de las opiniones políticas (principalmente centrándose estos en el electorado estadounidense, pues es el país que más ha potenciado este tipo de investigaciones), aunque casi todos estos estudios usan un enfoque basado en el modelo físico de intercambio cinético de los gases ideales, el cual asume que la probabilidad de que dos elementos de un sistema interactúen entre sí es la misma independientemente de las características de dichos elementos.

Por otra parte, las redes sociales van cobrando cada vez mayor protagonismo dentro de nuestras vidas, lo que ha motivado la realización de diversos estudios sobre los comportamientos humanos que derivan en la formación de dichas redes. Desde Twitter, Facebook, LinkedIn.... prácticamente todas presentan un modelo común, lo cual es coherente al tratarse de extensiones del comportamiento humano y de como este se desenvuelve en su día a día. Por ello, varios investigadores como Jure Leskovec o Paul Erdős han diseñados modelos de creación de grafos que tratan de emular la forma en la que los usuarios de dichas redes se comportan e interactúan entre sí de la forma más fiel posible.

De esta forma, surge la pregunta de como la opinión acerca de un determinado asunto puede variar incluyendo los condicionantes definidos por la topología de las propias redes sociales, es decir, una estructura en la que los usuarios únicamente interactúen con sus homónimos con los que se encuentran conectados, pues son aquellos cuya opinión puede resultar relevante e influir en la forma de pensar de cada uno. Para ello, aprovecharemos el modelo cinético mencionado en el segundo párrafo y lo adaptaremos a los trabajos de los estudiosos de las redes sociales para conseguir unos resultados más precisos.

1.2. Objetivos y estructura del trabajo

En este trabajo vamos a perseguir dos objetivos principales:

1.– El primero de todos es, tal y como se ha tratado en la sección anterior, diseñar un algoritmo que sea capaz de reproducir las simulaciones del modelo matemático de interacción de usuarios para la formación de opiniones que se describe en el modelo cinético de gases ideales adaptándolo a los condicionantes de la estructura de una red social. Para su consecución planteamos una serie de objetivos parciales:

1.1.– Estudiar el comportamiento de la opinión de la red social para una serie de diferentes estrategias de generación de grafos, poniendo especial énfasis en la planteada por Jure Leskovec.

1.2.– Estudiar el impacto que otras variables tienen en el comportamiento de la opinión de la población de la red de manera general, analizando en qué situaciones esta evoluciona desde el consenso hacia posiciones extremas.

2.– Diseñar una estrategia para, dentro de los resultados obtenidos en el objetivo anterior, incluir una serie de individuos que representan la función que tienen los medios de comunicación dentro de las redes sociales y observar como su presencia influye en la formación de las opiniones de los usuarios. Tras analizar las diferencias presentadas entre las dos partes, buscamos determinar si los medios de comunicación, con el modelo propuesto en este trabajo, influyen en la formación de la opinión en los usuarios de las redes sociales de manera general.

La estructura que esta memoria presenta coincide completamente con los objetivos que persigue. Cada uno de los objetivos principales viene desarrollado en capítulos diferenciados, mientras que cada objetivo parcial consta de una sección propia. De forma adicional, se introduce en el capítulo 2 conceptos técnicos básicos para el correcto entendimiento de la memoria y en el capítulo 5 las conclusiones finales tras realizar las simulaciones pertinentes y estudiar los resultados obtenidos.

Además, se presenta en el anexo A la explicación detallada y los resultados obtenidos en la realización de una encuesta popular que nos será de gran utilidad para el desarrollo de la sección 4.

ESTADO DEL ARTE

A lo largo del desarrollo de este capítulo, presentaremos en primer lugar los modelos de formación de opiniones más relevantes y más estudiados a lo largo de los últimos años. Continuaremos explicando las nociones básicas para ser capaz de comprender los distintos modelos de creación de grafos que serán empleados en el capítulo 3. En último lugar, se presentará de forma breve los principios de la función de distribución beta, que es ampliamente utilizada en el modelo propuesto en el capítulo 4.

2.1. Modelos de formación de opiniones

2.1.1. El modelo cinético de gases ideales

Los modelos de formación de opiniones que más relevancia han adquirido dentro de la comunidad científica son aquellos que siguen el modelo físico de intercambio cinético de los gases ideales que describe la forma en la que las partículas gaseosas dentro de un recipiente intercambian energía cinética [3]. Para comprender este modelo hay que recordar que un gas está formado por una serie de partículas que se distribuyen de forma uniforme dentro del recipiente que los contiene. Estas partículas, cada una con una energía cinética propia, chocan de manera aleatoria con el resto con igual probabilidad, confiriendo parte de su energía cinética.

Varios estudios que tratan esta transferencia de energía han sido realizados, al tratarse de un asunto que cuenta con aplicaciones, entre otras, en el mundo de la aeronáutica, al ser base de los principios aerodinámicos para el diseño de estructuras [17].

A raíz de esta serie de estudios surgió la idea de analizar el comportamiento la formación de opiniones usando un procedimiento similar. Así, cada una de las partículas se correspondería con un individuo, cada uno con una opinión propia acerca de un determinado asunto. Tras seguir las dinámicas dominadas por una ecuación matemática que modela cómo se produce el intercambio de opiniones, estos varían su opinión realizando sucesivas interacciones con el resto de individuos del sistema.

Existe una gran variedad de estudios que toman este modelo como base de sus investigaciones, variando principalmente en la definición de la ecuación que regula la interacción entre individuos. De

entre todos estos, vamos a centrarnos en el modelo propuesto por Mehdi Lallouache en [12], pues es el que usaremos como base para las simulaciones vistas en los capítulos 3 y 4.

El modelo de Mehdi Lallouache

En 2010, Mehdi Lallouache et al. [12], presentó un artículo en el que presentaba un modelo que mejoraba a otros modelos, presentando, entre otras novedades, un rango de opiniones continuo. En trabajos previos [7] se había propuesto un modelo estadístico multiagente en el que, al modelar la interacción de los individuos, entra en juego, además de las opiniones de los propios individuos, un parámetro ϵ ($0 \leq \epsilon \leq 1$ que se corresponde con una variable estocástica que varía con el tiempo y otro parámetro λ ($0 \leq \lambda \leq 1$) que podría interpretarse como la convicción o el sentimiento de apego de cada individuo hacia su propia opinión.

De esta forma, podemos estudiar el comportamiento del sistema cuando el número de individuos del mismo, $N(N \rightarrow \infty)$. Denotaremos la opinión del individuo i -ésimo de nuestro sistema en un tiempo t como $O_i(t)$, pudiendo esta adquirir valores entre -1 y 1 , determinando los dos extremos que pueden presentarse respecto de una opinión. Vamos a suponer que el tiempo es discreto, pudiendo este tomar únicamente valores naturales. Es denominado modelo multiagente homogéneo ya que asume un parámetro λ común para todos los individuos del sistema.

Definido de esta forma, la ecuación que modela la formación de cada individuo se describe como:

$$\begin{aligned} O_i(t+1) &= \lambda[O_i(t) + \epsilon_t O_j(t)] \\ O_j(t+1) &= \lambda[O_j(t) + \epsilon'_t O_i(t)]. \end{aligned} \tag{2.1}$$

El siguiente paso que parece pertinente es analizar el comportamiento de la opinión media de los sujetos a lo largo del tiempo y si esta, con valores elevados de t , se estabilizaba llegándose a un consenso entre todos los agentes o, por el contrario, se polarizaba. Para este análisis vamos a hacer uso de una variable que representa el orden de la opinión media en la sociedad tras un largo tiempo t ,

$$\bar{O}(t) = \sum_i^N O_i(t). \tag{2.2}$$

Tras realizar numerosas simulaciones, y analizando el comportamiento de $\bar{O}(t)$ frente al valor del parámetro λ , así como la variación de la varianza del primero frente a, de nuevo, λ , se llega a dos conclusiones relevantes. La primera es que en todas las simulaciones, el sistema evoluciona desde un estado de desorden aleatorio donde las opiniones están uniformemente distribuidas hacia un estado en el que o bien todas las opiniones se hayan polarizadas o se llega a un consenso. En segundo lugar se marca la existencia de un punto espontáneo de ruptura de simetría en el parámetro λ en el que el sistema se evoluciona desde el consenso a una situación de polarización en el que todas las opiniones

e encuentran en los extremos del rango de opiniones. Este punto, que se encuentra en $\lambda = 2/3$ posee la peculiaridad de presentar un comportamiento similar a los modelos de magnetización espontánea de los imanes de Ising [11] y es de especial relevancia al indicar que los sistemas, que en este caso intentan emular a las redes sociales, con un valor de convicción inferior a $2/3$ llegan a un consenso global a medida que el tiempo avanza mientras que presentan una importante polarización para valores mayores.

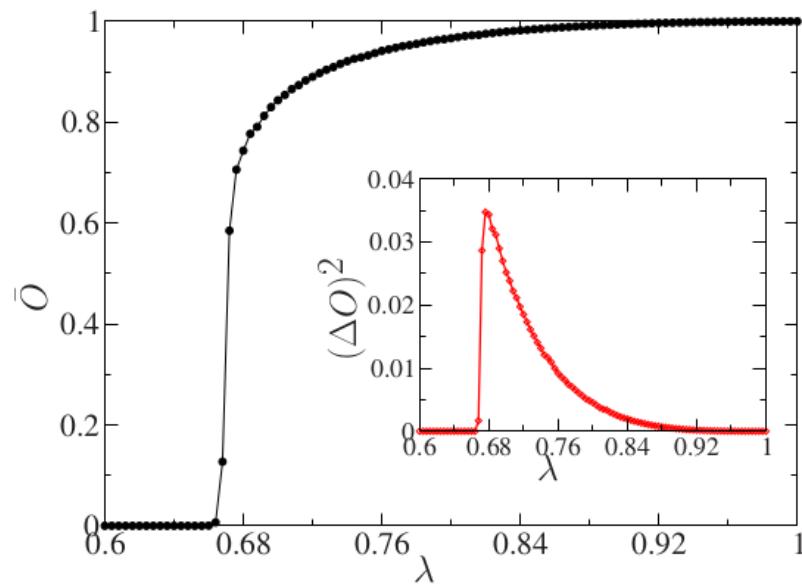


Figura 2.1: Resultados numéricos para las variaciones de la opinión media $\bar{O}(t)$ para t grandes con respecto de λ . (Recuadro) Resultados numéricos para la variación de la varianza $(\Delta O)^2 \equiv (O - \bar{O})^2$. Figura obtenida de la pág. 3 de [12].

2.1.2. El modelo del votante

El modelo del votante, fue introducido en la teoría matemática de probabilidad en 1975 y se trata en un modelo similar al que acabamos de ver, pero con la particularidad de que las opiniones de cada individuo alternan entre -1 y 1 a lo largo del tiempo en función de las influencias que los vecinos de dicho individuo pueden tener sobre su percepción acerca de un determinado asunto. Es un modelo de gran relevancia, pues es de los pocos que modela de manera precisa un sistema de varias partículas que interactúan entre sí. Además sirvió de punto inicial para describir una amplia variedad de fenómenos sociales [19].

Destaca de forma especialmente relevante el estudio de este modelo llevado a cabo por J. J. Ramasco y J. Fernández-García en 2014 [9], quienes analizan el grado de precisión de este y su correspondencia con la realidad, usando los datos públicos de las elecciones estadounidenses de 2000 y 2004. Estas elecciones son especialmente apropiadas para el modelo, pues debido a la ley electoral

de este país, en la que solo existen dos partidos en la carrera presidencial, democráticos y republicanos, presenta unas características óptimas a priori para el modelo del votante. Tras analizar la forma de la distribución que el modelo del votante arroja tras realizar las pertinentes simulaciones, Ramasco observa que este se corresponde en gran medida con los datos empíricos, tanto a nivel de condados como a otras escalas como distritos congresuales.

El modelo del votante puede presentar una serie de variaciones como el modelo no lineal del q-votante [6]. Tal y como lo hace S. Redner en su artículo *Reality-inspired voter models: A mini-review* [19] escrito en junio de 2019, el modelo del votante clásico puede definirse de la siguiente forma:

Sean N nodos, que se corresponde cada uno con un votante distinto, pertenecientes a un grafo estático, la evolución de la opinión viene definida por el siguiente algoritmo:

- 1.– Escogemos un votante de manera uniformemente aleatoria.
- 2.– Este votante adopta el estado de un vecino aleatorio.
- 3.– Se repiten los pasos 1 y 2 hasta que se alcance el consenso del sistema, entendiendo este como una situación en la que todos los nodos posean la misma opinión.

De forma matemática se puede expresar esta norma de actualización de la siguiente forma: sea $O_i = \pm 1$ el valor de la opinión del votante i -ésimo, el ratio al que dicho votante cambia de estado, w_i viene definido por la siguiente fórmula:

$$w_i = \frac{1}{2} \left(1 - \frac{O_i}{z} \sum_{j \in J_i} O_j \right), \quad (2.3)$$

donde z se define como el numero de coordinación de cada nodo y J_i denota el conjunto de vecinos del nodo i del grafo.

El modelo del votante, de nuevo, tiene numerosas características similares a las presentadas en el modelo de Ising [11] pese a que inicialmente, este se ideó como un modelo físico para el comportamiento de las partículas ferromagnéticas. El modelo de Ising ha tenido una gran importancia a lo largo de los años, debido principalmente a que fue uno de los primeros modelos que surgieron en el campo de la Mecánica Estadística, además de ser uno de los pocos que consta de una solución analítica exacta.

De forma resumida, el modelo de Ising plantea la existencia de un número N de partículas en un sistema, cada una con un espín, esto es, cada partícula posee una orientación, hacia arriba o hacia abajo. El modelo trata de determinar el sentido del espín a través de las interacciones de las partículas con sus vecinas y de una serie de fluctuaciones térmicas que se producen en el proceso. Formalmente, este modelo define el hamiltoniano del sistema de la siguiente manera:

$$\mathcal{H} = \sum_{\langle i,j \rangle} -J_{ij} \sigma_i \sigma_j \quad (2.4)$$

donde el hamiltoniano \mathcal{H} define la energía del sistema y el sumatorio abarca únicamente las partículas

$< i, j >$ que sean vecinas, siendo el espín de la partícula n -ésima determinada por $\sigma_n \in -1, 1$. El modelo de Ising contempla el parámetro J_{ij} como un factor de escala para representar las interacciones que se producen entre los espines de las partículas y la energía de dos partículas i y j vecinas.

2.1.3. Formación de opiniones con tolerancia limitada

No obstante, parece relevante, con el fin de acercarnos lo más posible a un comportamiento que pueda asemejarse al de la realidad, la introducción de un nuevo factor en el modelo presentado en el punto 2.1.1, que es el de la tolerancia. Tal y como realizan Rainer Hegelsmann y Ulrich Klause [18] en su artículo *Opinion dynamics and bounded confidence models, analysis, and simulation*, esto se consigue simulando una situación en la que, con un elevado número de nodos, cada uno con una opinión generada de forma simultánea, la opinión de estos únicamente se ve influenciada por aquellos nodos cuya opinión sea ϵ_r unidades mayor que la propia y ϵ_l unidades menor.

Este concepto puede implementarse de diferentes formas. La primera de ellas, mostrada en la figura 2.2, consiste en tomar valores $\epsilon = \epsilon_r = \epsilon_l$, obteniendo la evolución de las opiniones y los valores de ϵ para los cuales se llega a alcanzar estados de consenso entre los nodos del sistema. El principal resultado que se observa es que el sistema avanza desde la pluralidad para valores de ϵ menores a 0,15, para pasar a un intervalo en el que se polarizan las opiniones, para finalmente llegar a un consenso general para valores de ϵ mayores a 0,25.

Otra posible forma de implementar el concepto de tolerancia, puede ser a través de intervalos de confianza asimétricos: $\epsilon_l = \phi\epsilon_r$; $\phi \in (0, 1)$. En este caso los resultados que arroja el modelo permiten ver el estado final de la opinión media para cada par de valores como queda representado en la figura 2.3.

Como alternativa final, se puede implementar una versión, esquematizada en la figura 2.4 de intervalos de confianza asimétricos en los que la diferencia entre ϵ_r y ϵ_l varía en función de la distancia de la opinión de cada nodo a los extremos del intervalo $[1, -1]$. Los resultados obtenidos en este caso muestran que, para los casos en los que la diferencia entre intervalos es mayor, los extremos se ven afectados de menor manera por aquellas opiniones ubicadas en el centro, conllevando a una polarización de las opiniones hacia los extremos, lo que se corresponde con el resultado intuitivo de este experimento.

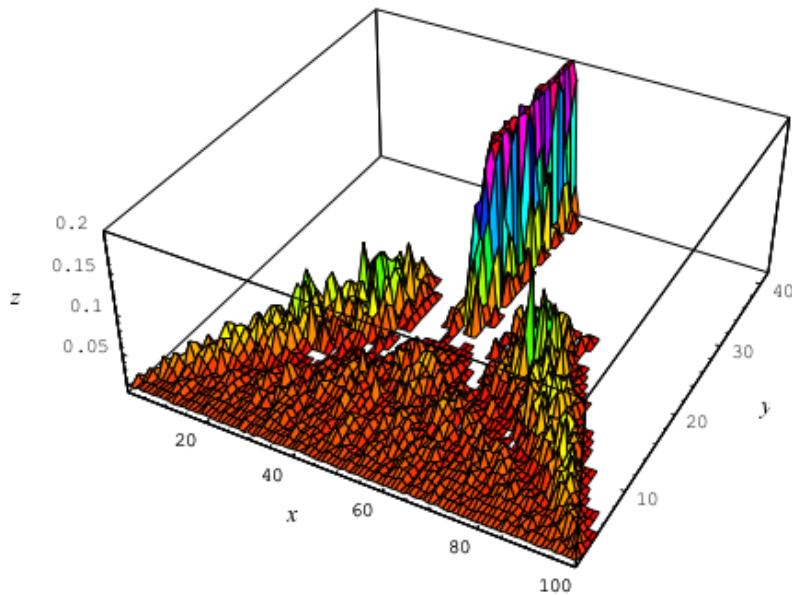


Figura 2.2: Simulación realizada con 625 nodos con opiniones distribuidas de forma uniforme tras realizar 15 períodos de interacción y con valores de ϵ simétricos. El eje x se corresponde con las 100 divisiones del rango de opiniones $[0, 1]$, el eje y representa el valor (multiplicado por 100) de ϵ y el eje z representa la frecuencia relativa de las opiniones en cada uno de los 100 intervalos. Figura obtenida de la pág. 11 de [18].

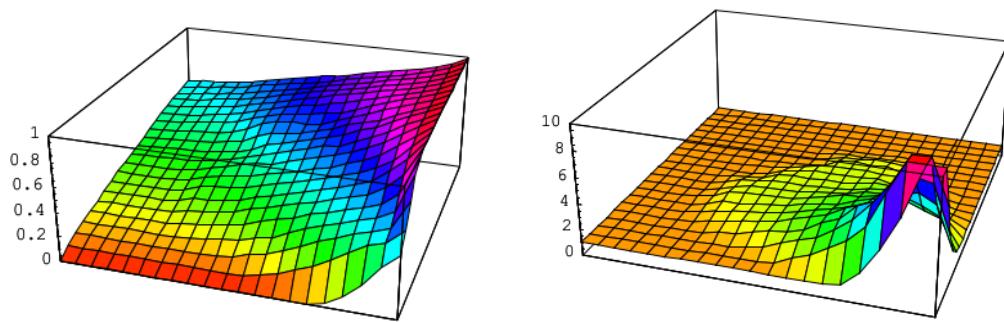


Figura 2.3: En ambos casos el eje x representa el valor de ϵ_l , el eje y el de ϵ_r , ambos comprendidos entre el 0 y el 0,4. De nuevo se trata de una simulación con 625 nodos y 25 iteraciones de intercambio de opinión. En la figura izquierda el eje z representa la opinión media de los nodos tras la estabilización del sistema, mientras que en la derecha representa el número de opiniones diferentes supervivientes tras la estabilización del sistema. Figura obtenida de las págs. 21 y 22 de [18].

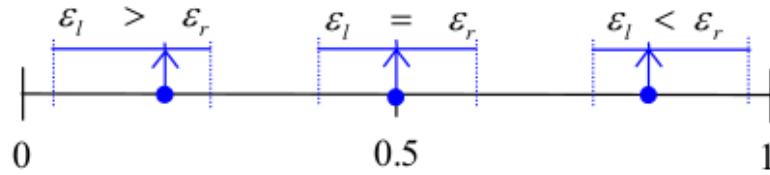


Figura 2.4: Variación de los intervalos de confianza de Hegelsmann. Se observa que a medida que las opiniones se hayan más cercanas de los extremos, el intervalo de confianza correspondiente con el extremo más cercano a dicho extremo crece mientras que el más alejado disminuye. Figura obtenida de la pág. 23 de [18].

2.1.4. Otros modelos de formación de opiniones

Otros modelos, en los que no entraremos tanto en detalle, han sido desarrollados a lo largo de los años por la comunidad científica.

Nos gustaría destacar la labor realizada por Giuseppe Toscani y Carlo Brugna [22] [5], en el que, usando unas notables habilidades de desarrollo matemático, estos investigadores italianos proponen un nuevo modelo de formación de opiniones basado en las interacciones binarias involucrando propiedades de difusión y de consenso, cada una representadas por parámetros distintos. El modelo que desarrollaron analiza un espectro continuo de opiniones ubicadas entre -1 y 1 y emplea ecuaciones diferenciales parciales a través de las cuales son capaces de llegar a la ecuación de Fokker-Planck. Esta ecuación, ampliamente usada en mecánica estadística para describir la evolución temporal de la función de densidad de probabilidad de la velocidad de una partícula bajo influencia de una serie de fuerzas, permite realizar un estudio de los momentos de la misma de forma cerrada.

También queríamos destacar el trabajo de Parongama Sen [20], en el que se usa como punto de partida el modelo desarrollado en la sección 2.1.1. La principal diferencia que introduce consiste de una nueva variable en la ecuación que afecta a la forma en la que la formación de la opinión propia se ve influenciada por el resto de agentes cuando la interacción binaria se produce. De esta forma, la ecuación que rige el modelo de formación de opiniones resulta:

$$\begin{aligned} O_i(t+1) &= \lambda_i O_i(t) + \epsilon_t \mu_j O_j(t) \\ O_j(t+1) &= \lambda_j O_j(t) + \epsilon'_t \mu_i O_i(t). \end{aligned} \tag{2.5}$$

En este caso, mientras que el parámetro λ representa el apego a la opinión propia y ϵ se trata de una variable aleatoria dependiente de la interacción en la que nos encontramos, μ se utiliza para plasmar la influencia que un agente concreto tiene sobre el resto de nodos, siendo este un parámetro fijo a lo largo de las interacciones.

2.2. Las redes sociales como grafos

Una red social consiste en una estructura compuesta por un conjunto de agentes y uno o más conexiones o relaciones definidas entre ellos. Han sido estudiadas desde principios de la década de 1930. En 1940 se produciría un importante avance en estos estudios, al incorporarse a estos la teoría de grafos para modelar el análisis de las redes sociales.

Un grafo es una pareja ordenada $G = (V, E)$ en la que el término V denota un conjunto no vacío de vértices o nodos mientras que E se refiere a un conjunto de aristas o conexiones definidas mediante pares no ordenados¹ de vértices $\{x, y\}$.

Al representar una red social con la estructura de un grafo, cada componente de dicha red se representarían por medios de los vértices y si dos vértices tuvieran la relación predeterminada para dicha red social, entonces presentarían una arista que les uniera. Dependiendo del diseño de la red social podríamos hablar de un grafo en el que la arista funcionara en un único sentido o en ambos, es decir, si existe una arista entre los vértices A y B , A se relaciona con B pero no viceversa. No obstante, a lo largo del desarrollo de esta memoria nos vamos a referir en exclusividad a los grafos no dirigidos, es decir, si A presenta una relación con B , esto significa que B presenta el mismo tipo de relación con A .

2.2.1. El modelo de Leskovec

En 2008 Jure Leskovec, junto a su equipo de Las Vegas, diseñó y llevó a cabo un estudio de la evolución de cuatro redes sociales online con información temporal sobre las creaciones de los nodos y de las aristas en el grafo que modelaba dichas redes [13]. Esta investigación ha sido de gran importancia, ya que se trata de la primera vez que se estudiaban los procesos de creación de aristas y de nuevos nodos a una larga escala, lo que permitió que el escrito adquiriera un enfoque macroscópico de las propiedades de las redes sociales.

Este trabajo permitió investigar una serie de estrategias de formación de redes basadas en principios de máxima-probabilidad, en las que la localidad de una arista juega un papel critico en la evolución de la red, y compararlas con los resultados que habían observado en el estudio previo para poder así diseñar un algoritmo que modelara de la forma más precisa posible la creación de un grafo que represente la topología definida por una red social. Para ello se basaron, no únicamente en la estructura global de la red estudiada, sino también en el comportamiento microscópico de cada nodo. Las cuatro redes sociales estudiadas fueron: Flickr, Delicious, Yahoo! y LinkedIn. Ambas comparten una estructura en la que los nodos representan personas mientras que las aristas se corresponden con relaciones sociales. A lo largo del proceso de recolección de datos, descompusieron su trabajo en tres procesos que se detallan a continuación:

¹En este caso se trata de pares no ordenados pues estamos tratando con grafos no dirigidos. Para la definición de un grafo dirigido deberíamos tratar con pares ordenados de vértices.

1.– El proceso de llegada de nodos nuevos a la red.

Observaron la existencia de variaciones importantes en las tasas de llegada de nuevos nodos a la red entre las cuatro estudiadas, diferenciándose estas incluso en el comportamiento de su crecimiento, por lo que decidieron tomar la tasa de llegada de nuevos nodos a la red como un parámetro de entrada dentro del propio modelo.

2.– El proceso de iniciación de aristas, que determina para cada nodo cuando establecer una relación social con otro a partir de su llegada a la red.

La conclusión final fue que las iniciaciones de aristas aceleran a medida que el tiempo pasa comprobaron que conlleva a unas distribuciones del orden de potencias.

3.– El proceso de selección de destino de la relación social.

En este proceso observaron que entre el 30 % y 60 % de las aristas formaban triángulos. Explicándolo con la estructura del grafo, esto significa que el destino de la arista se encuentra únicamente a dos saltos del origen y que principalmente los nodos tienden a crear aristas con otros nodos que se encuentran a distancias cortas.

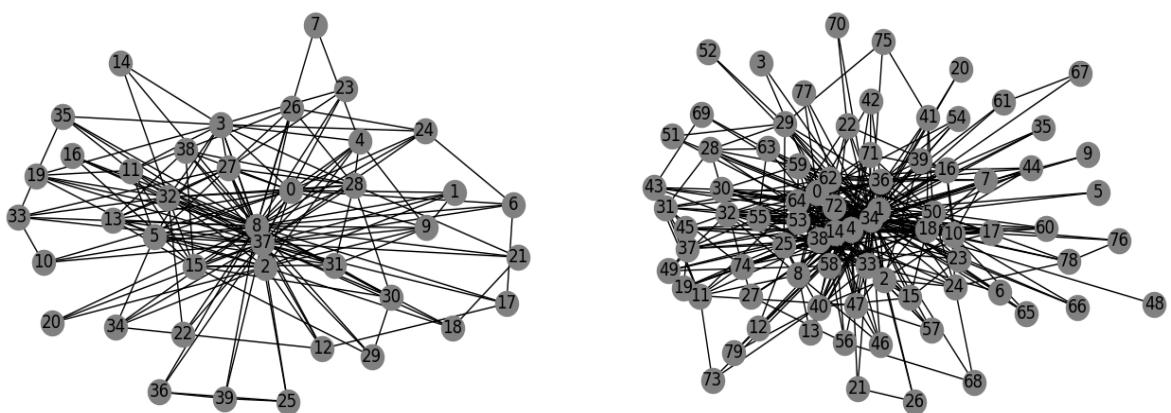


Figura 2.5: Ejemplos de grafos de Leskovec para redes con 40 y 80 nodos.

Con todo esto se puede diseñar un modelo de evolución de redes que puede ser usado para generar redes sintéticas de tamaño arbitrario que sean muy similares a las características macroscópicas de las redes sociales reales como puede observarse en la figura 2.5. Se trata del modelo de creación de grafos que hemos tomado como principal a lo largo del desarrollo de este trabajo, al haberse considerado que se trata del modelo más avanzado y que más fielmente representa la topología de las redes actuales.

2.2.2. El modelo de Erdös-Rényi

A continuación, vamos a presentar otro modelo de generación de grafos que es usado, de manera mucho más marginal, a lo largo del desarrollo de la sección 3.3. Este se trata del modelo de grafo de

Erdős–Rényi.

Este modelo, propuesto por Paul Erdős y Afréd Rényi en 1959 [8], toma de forma inicial una red compuesta por un total de N nodos distribuidos de forma aleatoria y sin ninguna arista entre ellos. El modelo prosigue conectando de forma aleatoria los nodos dos a dos un total de M veces por nodo. De esta forma se consigue un grafo con una conectividad media por nodo igual a M .

Hemos empleado este modelo ya que cobró especial importancia a partir de la década del 1950 y se usó para modelar redes complejas, sin embargo en los siguientes años este modelo ha sido reemplazado por otros que se ha comprobado que poseen una mayor precisión a la hora de representar una red social compleja.

2.2.3. El modelo de Barabási-Albert

El modelo de Barabási–Albert [1] fue un modelo popularizado en 1999 diseñado para representar sistemas aleatorios complejos sin tener en cuenta el tamaño de los mismos. Fue un modelo muy novedoso ya que incorporaba el mecanismo de conexión preferencial y mejoraba a modelos anterior como el presentado en la sección 2.2.2 ya que permitía representar más fielmente las redes que aparecen en la naturaleza pues sus distribuciones de grado (distribución del número de conexiones de un nodo de la red) siguen leyes de potencias tal y como se ve en las redes reales.

Para entender el modelo de Barabási-Albert, hay que tomar una red que parte con un número N de nodos conectados aleatoriamente de forma que todos los nodos tengan al menos una conexión. A continuación, cada nodo es conectado a un total de m nodos de forma proporcional al número de enlaces existentes en la propia red. De esta forma se consigue la existencia de una serie de nodos que reciben de manera marcada más enlaces que el resto, lo que podría considerarse como una serie de "influencers" dentro de la red social que tratamos de modelar. Formalmente podemos determinar la probabilidad de que un nuevo nodo se conecte a un nodo n -ésimo de la siguiente manera:

$$p_n = \frac{k_n}{\sum_i k_i}, \quad (2.6)$$

donde k_i denota el número de conexiones del nodo i -ésimo.

Este mecanismo se denomina conexión preferencial, pues los nuevos nodos de la red, que se van añadiendo uno a uno, poseen *preferencia*.^a ser enlazados con los nodos con mayor número de conexiones previas.

2.3. La función de distribución de probabilidad beta

A continuación hacemos un breve repaso sobre la función de distribución de probabilidades beta. Esta nos será de gran utilidad a la hora de desarrollar el capítulo 4, donde se usará como parte central de un algoritmo para predecir la forma en la que los usuarios de una red social deciden a qué medios de comunicación prestarán atención basándose en la distancia de la opinión de dichos medios a la propia.

La función de distribución beta es una familia de funciones de distribución continuas de probabilidad. Posee la peculiaridad de que toma valores en el intervalo $[0, 1]$ y ha sido ampliamente utilizada desde que Thomas Bayes la propusiera en un artículo póstumo publicado por Richard Price [2]². Esto se deba que permite ajustarse a una gran variedad de distribuciones empíricas, ya que su forma queda completamente determinada por dos parámetros positivos α y β . Estas variables permiten determinar la proporción de la integral de la función de densidad de probabilidad beta que caerá a la izquierda y a la derecha del punto medio; 0,5.

Formalmente diremos que una variable aleatoria continua X tiene una distribución beta con parámetros $\alpha > 0$, $\beta > 0$ (y diremos que $X \sim B(\alpha, \beta)$) si su función de densidad viene definida por:

$$f_X(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad (2.7)$$

para valores $0 < x < 1$ donde la función beta ($B(\alpha, \beta)$) se define de la siguiente forma:

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx. \quad (2.8)$$

Esta función de distribución tiene una serie de características muy notables, destacando que tanto su media como su varianza vienen determinadas únicamente por el valor de los parámetros α y β :

$$E[X] = \frac{\alpha}{\alpha + \beta} \quad (2.9)$$

$$Var(X) = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2} \quad (2.10)$$

Veremos la utilidad de la función de distribución beta, en concreto de su función de densidad de probabilidad en el capítulo 4. Nos interesa notar que para los casos en los que $\alpha = \beta$, la integral de la función de densidad de probabilidad beta será igual a ambos lados de $x = 0,5$. De igual forma vemos con las subfiguras (c) y (d) de la Figure 2.6 que la función puede cambiar su forma en función del valor de estos pese a que ambas variables sean iguales. Efectivamente, llamando $\alpha = \beta = \gamma$, vemos que la

²De hecho, la sociedad científica no empezó a utilizar la distribución beta hasta muchos años después, a raíz de los estudios del matemático inglés Karl Pearson [24], quien vivió entre finales del siglo XIX y principios del siglo XX.

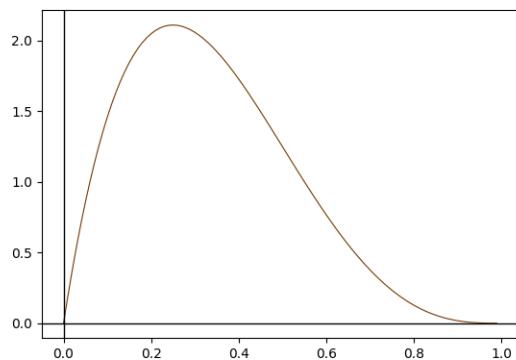
esperanza de la distribución:

$$E[X] = \frac{\alpha}{\alpha + \beta} = \frac{\gamma}{\gamma + \gamma} = \frac{1}{2}, \quad (2.11)$$

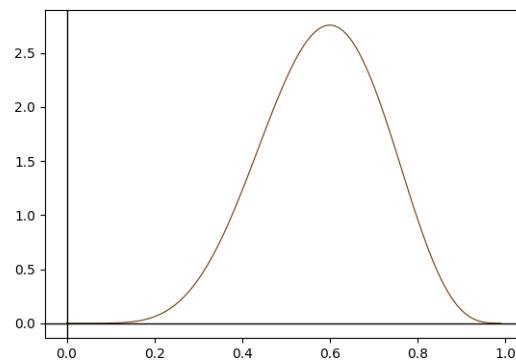
independientemente del valor de γ . Sin embargo, la varianza de la distribución

$$Var(X) = \frac{\alpha\beta}{(\alpha + \beta + 1)(\alpha + \beta)^2} = \frac{\gamma^2}{(2\gamma + 1)(2\gamma)^2} = \frac{1}{8\gamma + 4} \quad (2.12)$$

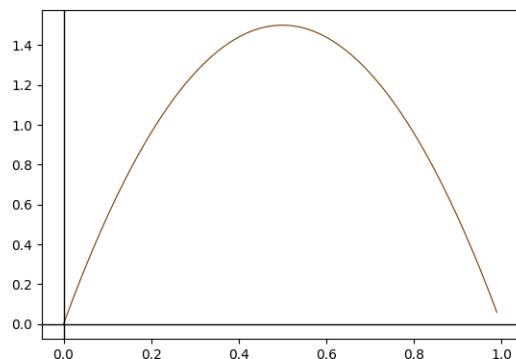
sí que depende del valor de γ , observándose un comportamiento descendente a medida que aumentamos el valor de este parámetro.



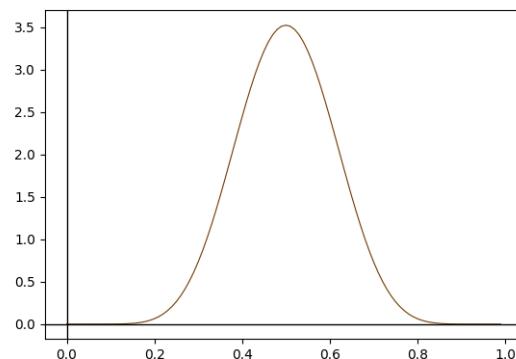
(a) $\alpha = 2, \beta = 4$



(b) $\alpha = 7, \beta = 5$



(c) $\alpha = 2, \beta = 2$



(d) $\alpha = 10, \beta = 10$

Figura 2.6: Funciones de densidad de probabilidad beta para diferentes parámetros α y β . Puede observarse en las dos primeras cómo el hecho de que uno de los parámetros sea mayor que el otro determina que la función se deforme hacia un lado o a otro del punto $x = 0,5$. En las dos últimas gráficas puede observarse también como la función de densidad de probabilidad, a medida que sus parámetros aumentan de valor, muestra unas pendientes mayores y, por lo tanto, unos máximos mayores.

EL MODELO CINÉTICO DE GASES IDEALES Y LAS REDES SOCIALES

En esta sección presentamos los resultados obtenidos al incorporar al modelo cinético de gases ideales la teoría de grafos para definir la topología de una red social, siendo nuestra parte central aquella que desarrolló Jure Leskovec en su artículo *Microscopic evolution of social networks* [13], permitiéndonos obtener una serie de resultados más precisos para analizar la evolución de la opinión media de los sujetos pertenecientes a una red social.

A lo largo del desarrollo de este capítulo, así como en el siguiente, debemos hacer una aclaración con respecto a lo que nos referiremos a la hora de mencionar ciertos conceptos. El fin último de nuestra investigación es el de hallar las condiciones en las que los individuos que componen la red social tienden a acumularse en torno a un determinado valor (generalmente en torno al punto medio del rango de opiniones) o, por el contrario, al finalizar las interacciones de las pertinentes simulaciones, los individuos tienden a adoptar opiniones extremas.

Por ello cuando expresemos que nuestra simulación ha alcanzado una situación de consenso, estaremos refiriéndonos al caso en que todos, o prácticamente todos, los individuos han finalizado la simulación teniendo una opinión similar, y veremos que esta se centra siempre en el 0. Sin embargo, diremos que los individuos de la red social se han posicionado de forma extrema cuando estos tengan opiniones muy cercanas a 1 o a -1, siendo indiferente que todos posean la misma opinión o estén repartidos en cualquier proporción entre estos valores extremos.

Debido a esto, cuando representemos la opinión media de las simulaciones, lo que realmente estaremos haciendo es representar la media del valor absoluto de las opiniones de cada uno de los nodos del grafo. De esta forma trabajaremos con gráficas mucho más reveladoras, ya que eliminaremos las interferencias producidas en las simulaciones, pues todas ellas tienen componentes de aleatoriedad que les hacen oscilar entre el 1 y el -1 en numerosas ocasiones.

3.1. El modelo de Leskovec y la evolución de la opinión media

El primer paso de la investigación presentada en este trabajo comienza con el artículo de Lallouache comentado la sección 2.1.1. Este artículo aporta una serie de resultados reveladores en torno a un valor de la variable λ para el cuál la forma de distribuirse de la media de las opiniones cambia drásticamente. Sin embargo, y como los propios autores comentan al final del escrito, existen diversas formas de mejorar el modelo para que este presente unos comportamientos más similares a lo que podemos observar dentro de la vida real.

El modelo cinético de gases ideales presenta un gran inconveniente a la hora de perseguir este objetivo. Dentro de una red social, un individuo no es influenciado por un usuario con el que no posea ningún tipo de interacción. Facebook tiene 2900 millones de usuarios¹ mientras que el usuario medio de esta plataforma no dispone más que entre 300 y 400 amigos de media². Parece ilógico pensar que cada uno de esos 2900 millones de usuarios pueden influenciar la opinión de cada usuario de Facebook, encargándose de ello las amistades que cada usuario mantiene dentro de la red social. Es por ello que hemos decidido incorporar al modelo matemático que permite calcular la evolución de la opinión de cada individuo presentado por Lallouache, el estudio de Leskovec mostrado en la sección 2.2.1 que permite representar la estructura de una red social mediante una estrategia de formación de grafos. Esta decisión coincide con lo que, a priori, puede resultar intuitivo, pues parece claro que la opinión de un individuo perteneciente a una cierta red social no se verá afectada de ninguna manera por la opinión de otro sin ningún tipo de característica común ni con lazos que le conecte dentro de dicha red.

Tras incorporar este modelo, en primer lugar estudiaremos la evolución de la opinión media a medida que sucesivas iteraciones se realizan en la ecuación que da forma a la evolución de opiniones para determinados valores de la variable λ con una red compuesta por 600 individuos. Se ha planteado un sistema en el que dos individuos interaccionan entre sí de manera simultánea únicamente si son vecinos (es decir, que disponen de una arista que les conecte dentro del grafo), actualizando su opinión en el momento en el que finaliza la interacción. Formalmente se puede definir la evolución de la opinión provocada por la interacción entre dos individuos i y j si existe una arista que los conecte de la siguiente forma:

$$\begin{aligned} O_i(t+1) &= \lambda(O_i(t) + \epsilon_i O_j(t)) \\ O_j(t+1) &= \lambda(O_j(t) + \epsilon_j O_i(t)), \end{aligned}$$

¹<https://es.statista.com/estadisticas/600712/ranking-mundial-de-redes-sociales-por-numero-de-usuarios/>

²https://www.eldiario.es/comoyporque/no-puedes-150-amigos-si-puedes_1_8720697.html

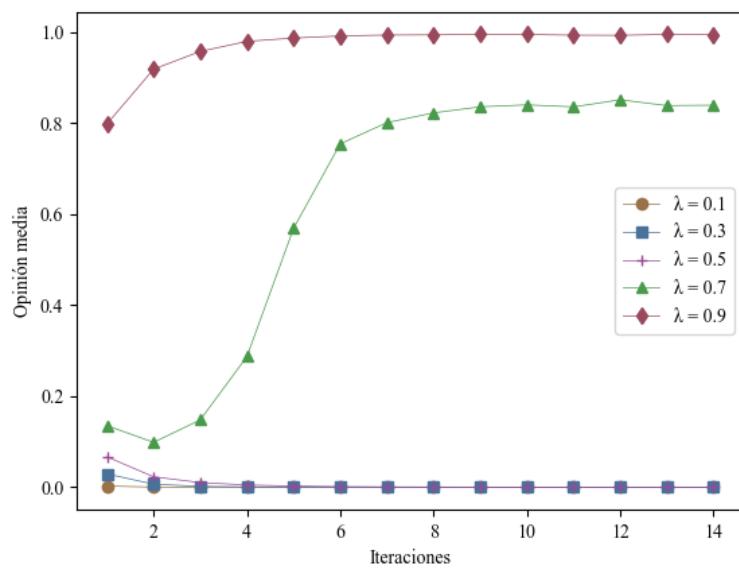


Figura 3.1: Comportamiento de la opinión media para distintos valores de λ . Se puede observar un comportamiento muy similar para aquellos valores $\lambda < 0.5$, mientras que cambia radicalmente para aquellos valores mayores.

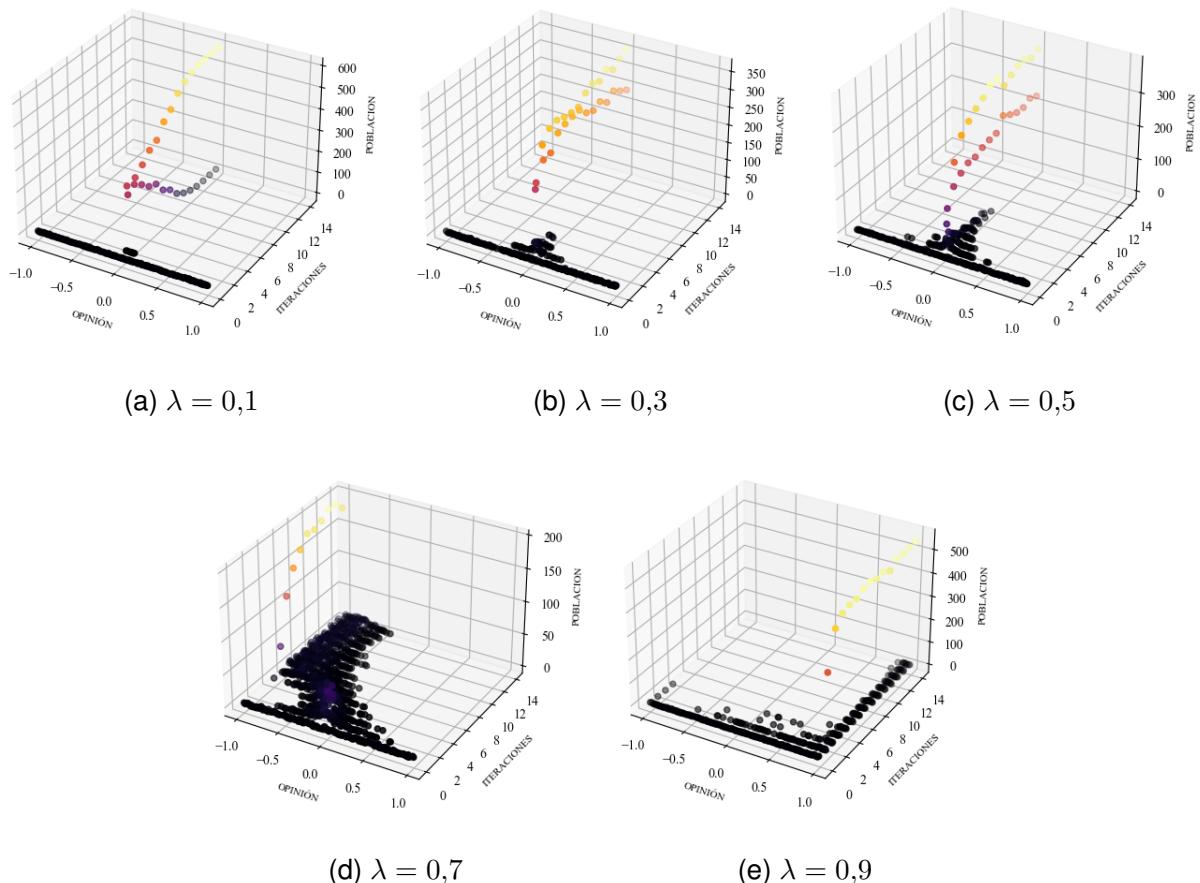


Figura 3.2: Distribución de las opiniones según avanzan las iteraciones para distintos valores de λ .

donde ϵ_i y ϵ_j son variables distribuidas aleatoriamente entre 0 y 1 para cada instante t y cada par de individuos. En este caso, la variable t se refiere al número de la iteración producida. Para cada iteración se realiza una interacción por cada una de las aristas de las que dispone el grafo. Se han calculado únicamente valores de $t < 15$, pues se ha observado que, para valores mayores, el comportamiento es similar con pequeñas variaciones que no afectan a la conclusión del modelo.

En las figuras 3.1 y 3.2 se muestran los resultados obtenidos para distintos valores de λ . Podemos observar que en los cinco casos, la totalidad de los nodos acaban convergiendo a una opinión, siendo en los tres primeros esta a una opinión muy cercana al 0 (se logra, por lo tanto, un consenso general) mientras que en los dos últimos converge a uno de los extremos. En estos casos vemos que, a mayor valor de λ , más rápido tiende el valor de la opinión media al extremo correspondiente.

Tras realizar diversas simulaciones observamos que, en estos dos últimos casos en los que el valor de $\lambda = 0,7$ y $\lambda = 0,9$, el valor al cual convergen las opiniones de los nodos puede ser tanto 1 como -1, no siguiendo ningún patrón. Esto es debido a los mínimos desordenes producidos por la aleatoriedad de la distribución de opiniones iniciales y de creación del grafo que se ven maximizados a medida que las sucesivas iteraciones se suceden.

```

input : Un grafo en una determinada iteración t
output: El mismo grafo en la iteración t + 1

1 def interacción(G):
2     for nodo in G: do
3         for nodoaux in GetNodosAdyacentes (nodo): do
4             if abs (nodo.opinion - nodoaux.opinion) < TOLERANCIA: then
5                 epsilon = uniform (0,1)
6                 nodo.opinion = LAMBDA * (nodo.opinion + epsilon*nodoaux.opinion)
7                 epsilon = uniform (0,1)
8                 nodoaux.opinion = LAMBDA * (nodoaux.opinion + epsilon*nodo.opinion)
9             end
10        end
11    end

```

Algoritmo 3.1: Algoritmo que define la interacción entre dos nodos adyacentes dado un valor al parámetro de tolerancia.

Con estos resultados, y siguiendo la idea del trabajo de Hegelsman y Klaus de la sección 2.1.3, se realizan una serie de simulaciones, mostradas en la figura 3.3, añadiendo el concepto de la tolerancia, en este caso de manera simétrica. Para estas simulaciones hemos establecido que λ adopte valores de 0,6 y 0,8 para cada caso. Para su programación simplemente hemos introducido un condicional en el que, antes de hacer la actualización de opiniones pertinente, se comprueba que la distancia entre

las opiniones de ambos individuos sea menor al umbral de tolerancia determinado previamente. El pseudocódigo se corresponde con el mostrado en Algoritmo 3.1.

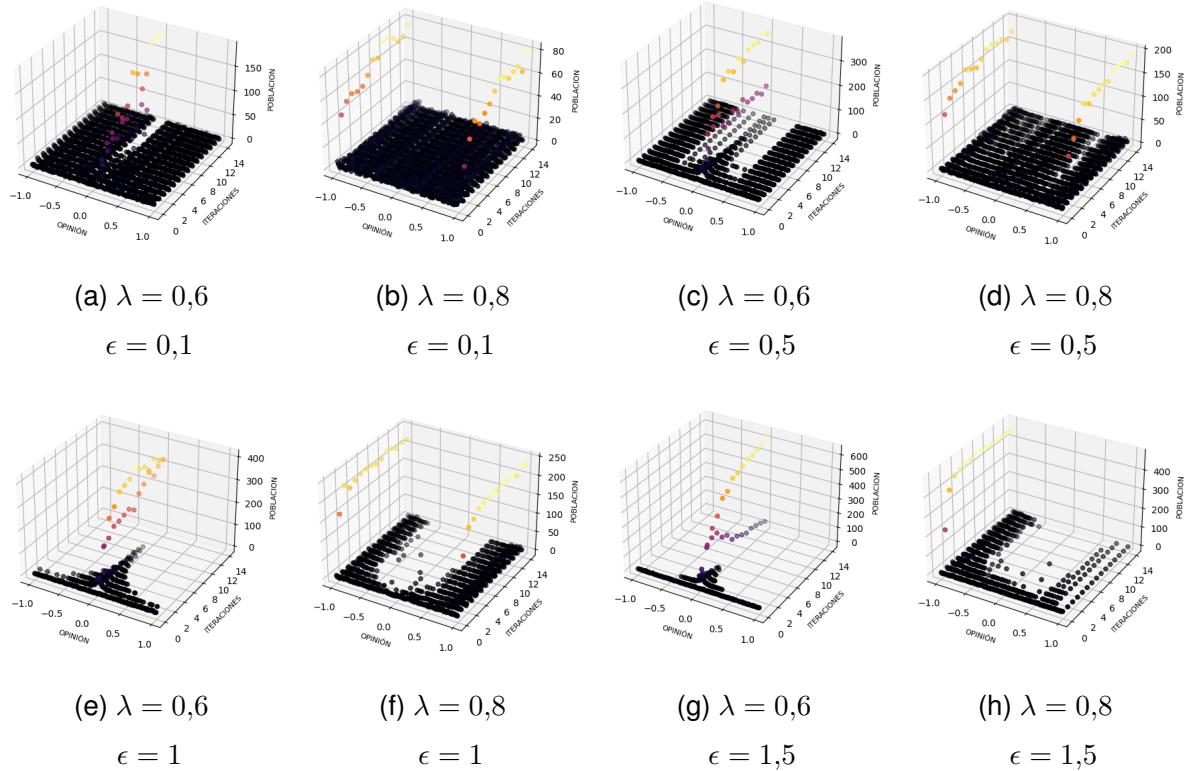


Figura 3.3: Distribución de la opinión a lo largo de la simulación con distintos valores de tolerancia. Se han realizado para cada caso simulaciones con $\lambda = 0,6$ y $\lambda = 0,8$.

Comprobamos en la figura 3.3 que, para los valores de tolerancia (ϵ) pequeños, se observa un alto número de individuos que no tienden en ningún momento ni al consenso ni a adoptar posiciones extremas, sino que mantienen su opinión de forma permanente a lo largo de las iteraciones que se realizan. Este resultado es fruto de que, debido a la topología del grafo, en el momento de la actualización de la opinión de un nodo, al solo depender de los vecinos de este, pueden producirse fácilmente *islas de opiniones* en las que un nodo esté únicamente conectado a otros cuya opinión se mantenga constantemente a una distancia mayor a la determinada por el umbral de tolerancia.

Sin embargo, la característica más notable es la mostrada en la figura 3.4. Observamos que, independientemente del valor del umbral de tolerancia, para todos los casos (recordemos que cada simulación ha sido realizada con grafos diferentes cada vez, lo cual introduce una mayor abstracción a los resultados) en los que el valor de $\lambda = 0,6$, en el sistema se observa una mayor tendencia al consenso en el que todos, o la mayoría de individuos poseen opiniones muy cercanas al 0, mientras que para los casos en los que $\lambda = 0,8$, la estructura general del sistema tiende bien a polarizarse en los casos de ϵ pequeños, bien a tender a que la mayoría de los nodos se decanten por una de las opiniones extremas tal y como sucedía en el caso anterior en el que los nodos interactuaban con sus

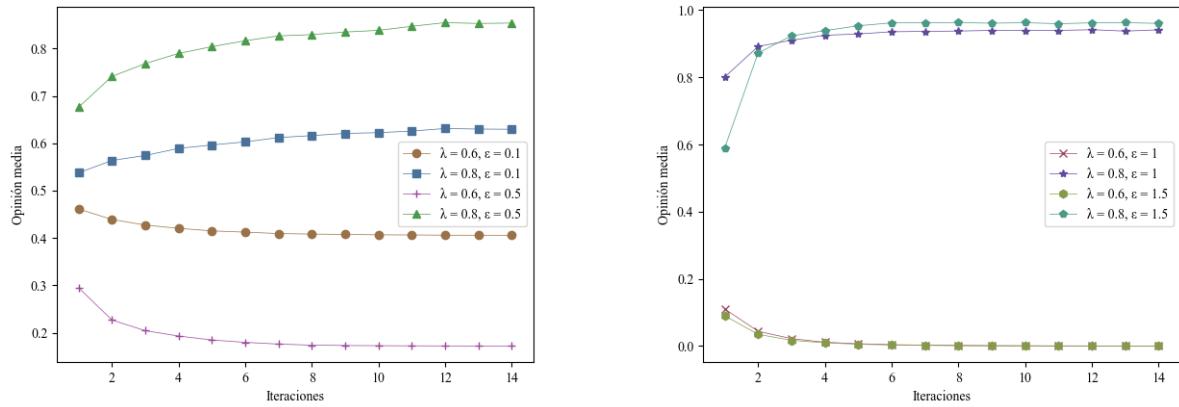


Figura 3.4: Evolución de la opinión media a medida que avanzan las iteraciones para distintos valores de tolerancia con $\lambda = 0,6$ y $\lambda = 0,8$. Podemos observar que se empiezan a reconocer ciertos patrones cuando para los valores de λ independientemente del valor de ϵ .

vecinos independientemente de la diferencia entre las opiniones de estos

3.2. Estudio de la evolución de la opinión en función del parámetro λ

Motivada por estos últimos resultados, surge la siguiente pregunta *¿Existe algún valor de λ para el cuál la opinión media del sistema transita desde el consenso hacia el posicionamiento en opiniones extremas?*.

Para responder a esta pregunta, se han diseñado una serie de simulaciones, cuyos resultados aparecen reflejados en la figura 3.5, en las que se ha decidido prescindir de la variable de la tolerancia, es decir, que los nodos, para actualizar su opinión, se ven influenciado por la opinión de todos sus vecinos, independientemente de la distancia de las opiniones de ambos ya que se ha observado. Esta decisión ha sido fruto de la realización de numerosas simulaciones con diferentes valores para esta variable, notando que no se trata de un parámetro relevante en la búsqueda de la respuesta a la pregunta planteada, además de añadir una nueva dimensión a nuestra simulación, lo cual complica enormemente tanto la propia simulación como la interpretación de los resultados obtenidos, no aportando ninguna información útil.

Observando la figura 3.5 puede apreciarse claramente la existencia de un valor, $\lambda = 0,65$ a partir del cual el valor de la opinión media comienza a aumentar rápidamente hasta que alcanzar un pico de, prácticamente 1, donde se mantendrá para valores de λ mayores. Echando un vistazo a la gráfica derecha, comprobamos que en este último caso, esto se corresponde con la localización de elementos a ambos extremos del espectro de opiniones, por lo que el sistema permanece polarizado de cualquier

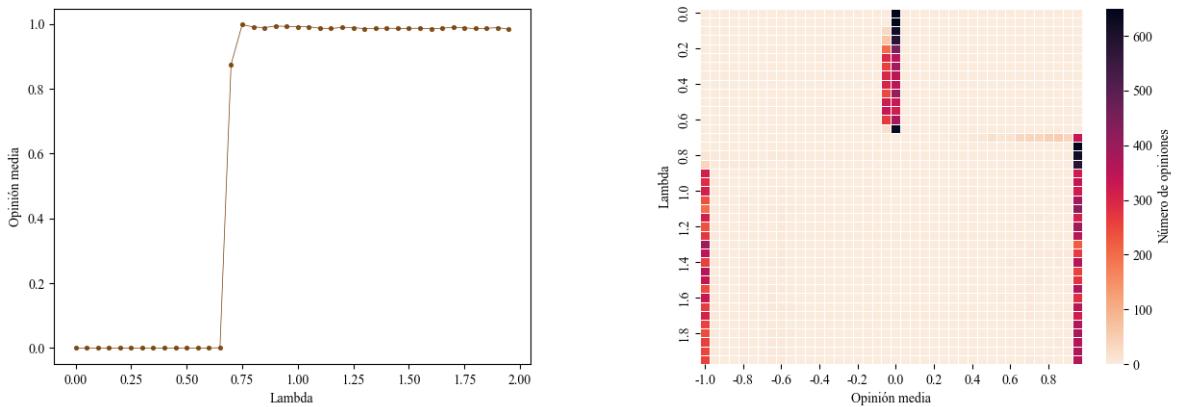


Figura 3.5: Comportamiento de la opinión media y de la distribución de las opiniones al aplicar la estructura de un grafo de Leskovec al modelo de formación de opiniones.

forma.

Se podría pensar que este resultado se obtiene para el caso específico del grafo de Leskovec generado en esta simulación, es por ello que se ha repetido este proceso un total de 25 veces y se ha tomado el valor de la opinión media, así como la varianza en la distribución de opiniones frente al valor de la λ , obteniéndose los resultados mostrados en la figura 3.6.

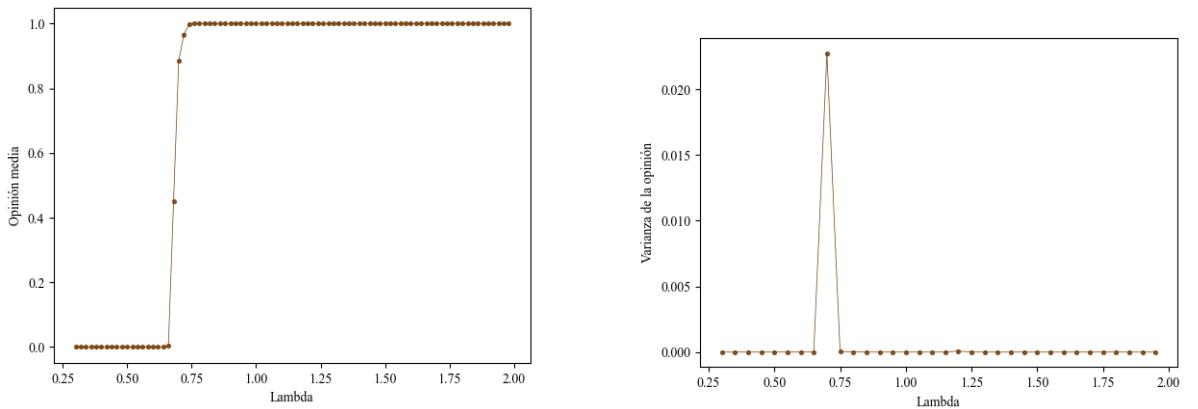


Figura 3.6: Comportamiento de la evolución media y de la varianza de la opinión respecto del valor de λ .

Vemos que, en efecto, para los valores en torno a $\lambda = 0,65$, la opinión media empieza a dispararse hasta alcanzar, un máximo de 1, lo que se corresponde con un valor de completa polarización en la que o bien todos los individuos de la red presentan una misma opinión extrema o están repartidos a ambos extremos del rango de opiniones. Analizando la varianza de las opiniones, vemos que esta es máxima para el valor $\lambda = 0,68$, lo que quiere decir que este es el valor para el cual mayor distribución

de opiniones dentro del espectro contemplado existe. En este caso, este punto, $\lambda = 0,68$, coincide con el mayor aumento en el valor de la opinión media. A partir de ahora, llamaremos a este punto, que se corresponde con el valor de λ a partir del cual se produce una transición desde el estado de consenso hasta el posicionamiento de las opiniones en los extremos del rango, como λ_c .

3.3. Otros modelos de creación de grafos

Tras obtener estos resultados, decidimos preguntarnos si este resultado se debe únicamente a la topología que presentan los grafos generados con la estrategia diseñada por Leskovec o, por el contrario, se produce en todo tipo de modelo de creación de redes sociales, por lo que decidimos probar con las siguientes estrategias de creación de grafos:

Grafo de Erdös-Rényi

Tal y como hemos explicado en la sección 2.2.2, el modelo de Erdös-Rényi se fundamenta en que el grafo posea un número medio de conexiones por nodo igual a un determinado valor m . Para determinar dicho m , hemos decidido crear un grafo de Leskovec con el mismo número de nodos y, posteriormente, calcular el valor del grado medio de la red:

$$m = \frac{\sum_{i=1}^{600} k_i}{600}, \quad (3.1)$$

siendo k_i el número de conexiones del nodo i -ésimo.

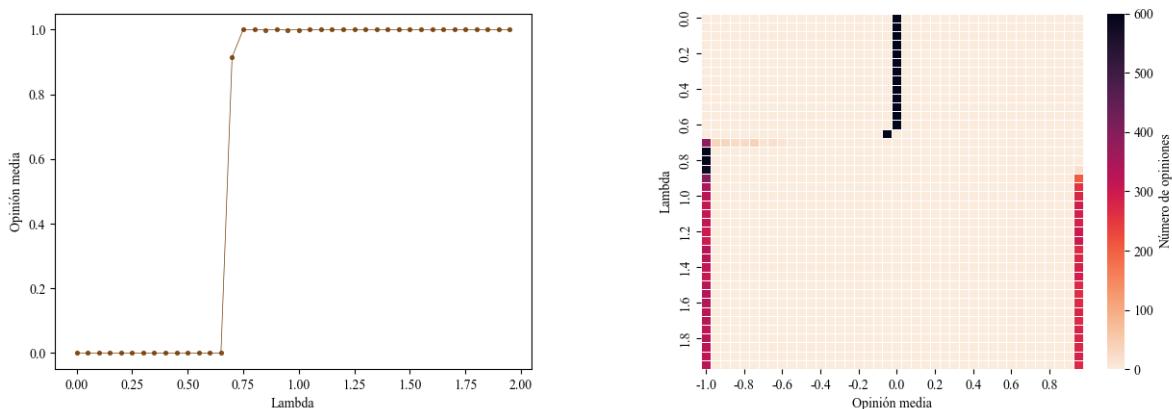


Figura 3.7: Comportamiento de la opinión media y de la distribución de las opiniones al aplicar la estructura de un grafo de Erdős-Rényi al modelo de formación de opiniones.

Tal y como se muestra en la figura 3.7, observamos que los resultados obtenidos muestran comportamientos muy similares a los obtenidos en la sección 3.2, produciéndose de nuevo un salto desde el consenso hasta el posicionamiento en opiniones extremas por parte de las opiniones de la red social

para valores de $\lambda = 0,67$, valor muy similar al obtenido en la anterior sección.

Grafo de Barabási-Albert

Como se ha visto en la sección 2.2.3, el modelo de creación de grafos propuesto por Barabási y Albert se basa en la conexión preferencial. Nuevamente, y tal y como se puede apreciar en la figura 3.8, vemos que, tras realizar las pertinentes simulaciones, el valor obtenido es, de nuevo, cercano a $\lambda_c = 0,68$. Para valores mayores de este valor, la opinión media de la población pasa de presentar un consenso global hasta llegar a posiciones de opinión extremas, al igual que en los apartados anteriores.

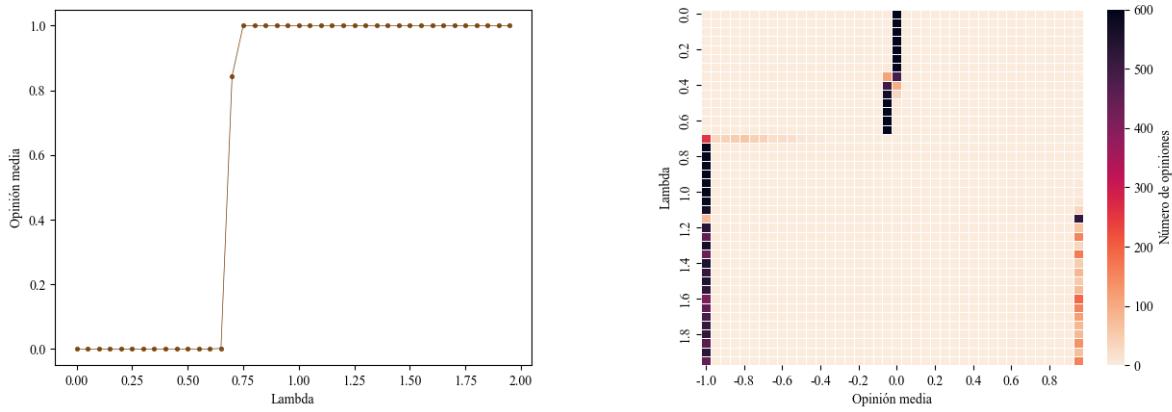


Figura 3.8: Comportamiento de la opinión media y de la distribución de las opiniones al aplicar la estructura de un grafo de Barabási-Albert al modelo de formación de opiniones.

Grafo completamente conectado

En último lugar, y a modo de experimentación, probamos con un grafo completamente conectado, esto es, en el que todos los nodos estén conectados con cada uno del resto de nodos del grafo. Este modelo, pese a que evidentemente no representa a ninguna red social (ni, probablemente, ninguna red creada por la naturaleza), presenta de nuevo las mismas características obtenidas en los modelos previos, obteniéndose un valor de $\lambda_c = 0,66$ tal y como puede verse en la figura 3.9.

Llegamos por lo tanto a la conclusión de que el valor crítico a partir del cual la opinión de la población perteneciente a una red social sufre un repentino cambio desde el consenso hasta la polarización, adopta valores similares para todas las estrategias de formación de grafos que hemos probado. Este resultado, que en todo caso se haya entre $\lambda_c = 0,66$ y $\lambda_c = 0,7$, coincide también con el mostrado en el estudio realizado por Lallouache [12], en el que determinaba $\lambda_c = 0,68$.

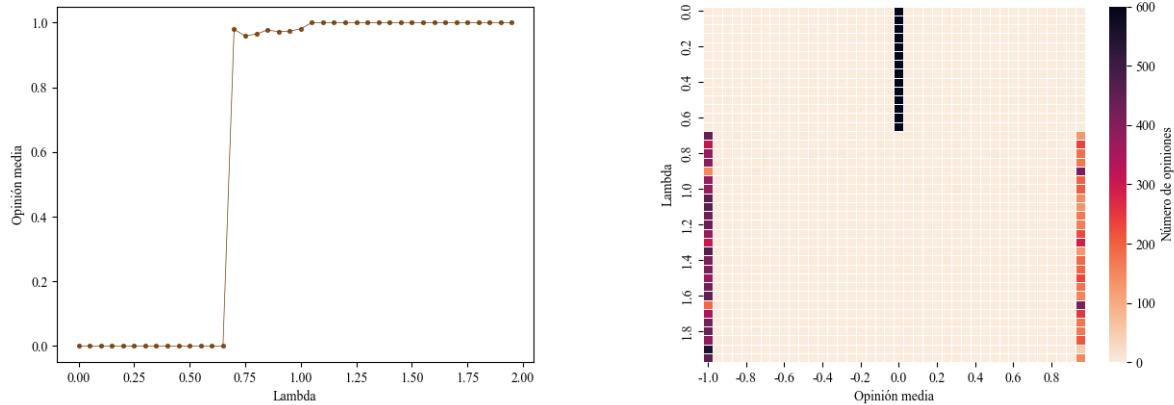


Figura 3.9: Comportamiento de la opinión media y de la distribución de las opiniones al aplicar la estructura de un grafo completamente conexo al modelo de formación de opiniones.

3.4. Otras consideraciones

En ultimo lugar vamos a analizar la influencia que una serie de condicionantes no planteados hasta ahora tienen en la determinación del valor de la λ a partir del cual se produce la transición en las opiniones de los componentes de una red social. Estos son:

Tolerancia

Ya hemos visto a final de la sección 3.1 que al conferir distintos valores al parámetro ϵ , que determinaba la distancia mínima respecto de la opinión propia para que se produjera la interacción, el valor de la λ_c no varía, sin embargo, sí que se observa una diferencia en el valor de la opinión media máxima a la que el sistema tiende para valores de $\lambda > 0,68$. Por ello en la figura 3.10 (a) representamos el valor máximo de la opinión media respecto del valor de la tolerancia.

Conectividad

De igual forma, y apoyándonos en el modelo de creación del grafo de Erdös-Rényi visto en el punto 2.2.2 que se basaba en un valor m que fija la conectividad media de cada nodo del grafo, mostramos en la figura 3.10 (b) la influencia del valor de m respecto del valor de λ_c .

Vemos que el valor de λ_c oscila únicamente de forma muy ligera, pues el máximo presentado es $\lambda_c = 0,7$ mientras que el mínimo $\lambda_c = 0,67$. Por esto, podemos afirmar que la conectividad no se trata de un condicionante que influya en la obtención del valor de λ_c .

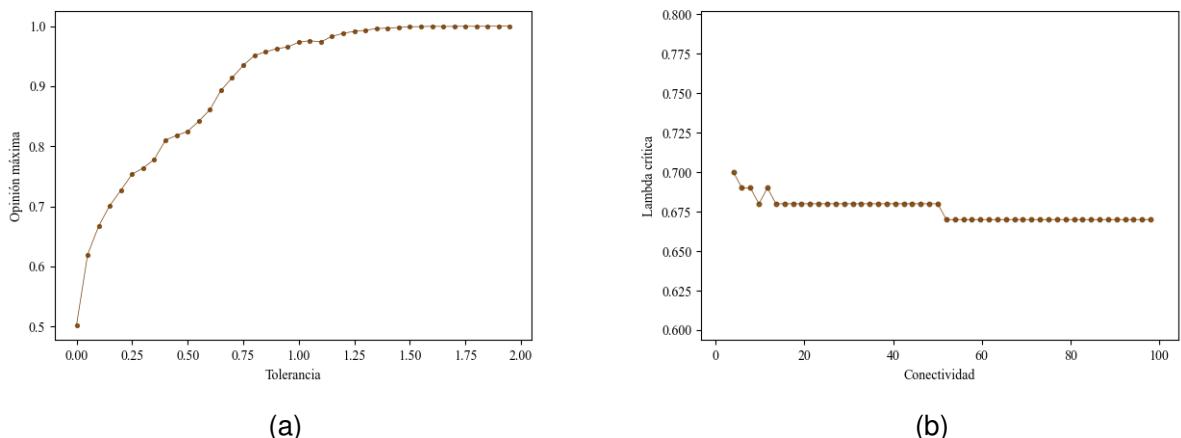


Figura 3.10: (a) Variación de la opinión media máxima en función del valor de la tolerancia. Observamos que hasta que se alcanzan valores de tolerancia cercanos a 1, la primera no llega a los valores cercanos a 1. (b) Evolución del valor de λ_c a medida que varía la conectividad media del grafo.

Número de nodos

Igualmente planteamos la opción de que el número de nodos de un grafo influya en los resultados obtenidos. Tras realizar simulaciones basadas en grafos de Leskovec de entre 50 y 1000 nodos, podemos ver que, nuevamente, el valor de λ crítica es independiente del número de nodos del grafo tal y como se muestra en la figura 3.11.

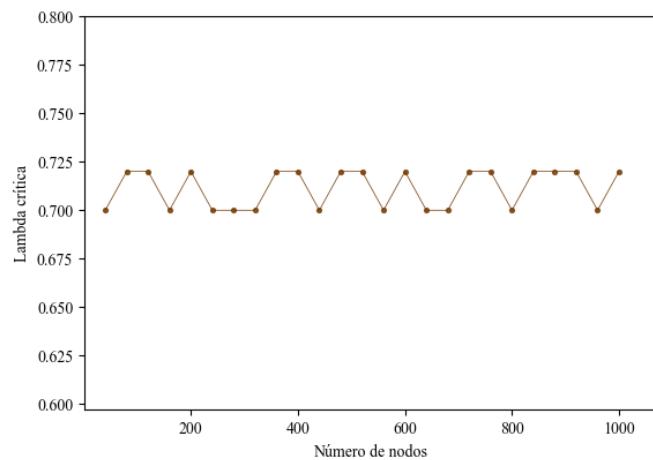


Figura 3.11: Variación del valor de λ_c para valores de nodos entre 40 y 1000. Observamos que, efectivamente, no se produce ninguna variación notable, pues la oscilación se produce entre $\lambda_c = 0,72$ y $\lambda_c = 0,7$, lo cual es una variación mínima.

Valor mínimo de ϵ en la ecuación 2.1

Para analizar la influencia del último parámetro en los resultados obtenidos, vamos a recordar la ecuación desarrollada por Lallouache (2.1):

$$\begin{aligned} O_i(t+1) &= \lambda(O_i(t) + \epsilon_i O_j(t)) \\ O_j(t+1) &= \lambda(O_j(t) + \epsilon_j O_i(t)), \end{aligned}$$

De manera predeterminada, y por continuar la línea de trabajo detallada en la sección 2.1.1, hemos considerado valores de ϵ comprendidos entre 0 y 1. No obstante, surge la duda sobre si esta variable, que en el fondo no trata de representar más que la influencia que tiene cierto agente sobre la opinión de cada nodo en el momento de las interacciones binarias, podría alcanzar valores comprendidos entre el -1 y el 1. Este límite se considera ya que, los valores inferiores al 0 se corresponderían con el hecho de que la opinión de cierto individuo causara rechazo y por ello sería recibida de manera negativa, intentando alejarnos de ella.

Por ello realizamos una serie de simulaciones mostradas en la figura 3.12, en las que representamos el valor de λ_c frente al valor mínimo de ϵ . Vemos que en este caso, el resultado, al contrario que en las 3 consideraciones anteriores, no es uniforme, apreciándose notables diferencias entre los valores de partida ($\epsilon = -1$) y los últimos ($\epsilon = 0$). De hecho, es posible describir un comportamiento general, pues el valor de λ_c desciende a medida que el valor mínimo de ϵ se acerca más al 0.

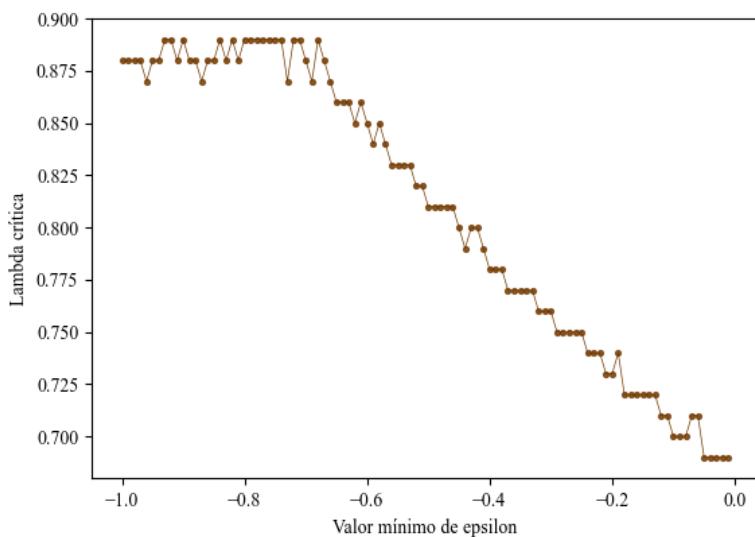


Figura 3.12: Variación del valor de λ_c a medida que varía el valor mínimo de ϵ en la ecuación de formación de opiniones. Vemos que a medida que ϵ adopta valores más cercanos al -1, el valor de λ_c aumenta.

Este se trata de un resultado notable, ya que, como hemos explicado, tanto λ representa el apego

a la opinión propia como ϵ el poder de influencia de un agente externo sobre la evolución de la opinión del propio nodo.

De este modo vemos que, si se pueden dar situaciones en las cuales un sujeto actúe de fuerza repelente, haciendo que nos alejemos de su opinión por el mero hecho de pertenecer a dicho individuo, será necesario un mayor sentimiento de convicción de manera general (recordemos que estamos tratando con poblaciones en las que todos los individuos presenten el mismo valor de λ) para que la red transite desde el consenso hasta situaciones en las que las opiniones de los componentes de la red social se sitúen en posiciones extremas.

En la figura 3.13, y a modo de comparativa, se muestran los resultados obtenidos tras realizarse las simulaciones presentadas en la sección 3.2, pero en este caso permitiendo a la variable ϵ alcanzar valores comprendidos entre el -1 y el 1. Vemos que, en efecto, se puede notar una notable diferencia con respecto a los mostrados en 3.5. En primer lugar podemos observar que ni el aumento vertiginoso que sufría el valor de la opinión media cuando $\lambda = 0,68$ se produce de una forma tan pronunciada, ni se produce en los mismos valores, siendo esta vez, el valor $\lambda_c = 0,88$ en el que se produce el mayor incremento.

Podemos notar también que, en esta ocasión, la evolución de todas las opiniones de los individuos de la red social desde el consenso en opiniones cercanas al 0 hasta localizarse en posiciones de opinión extremas, se produce de una manera mas lenta.

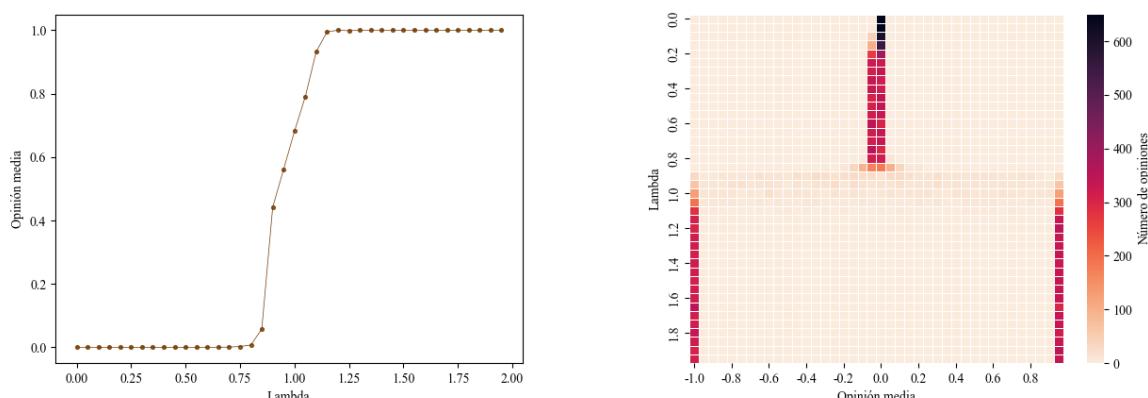


Figura 3.13: Comportamiento de la opinión media y de la distribución de las opiniones al aplicar la estructura de un grafo de Leskovec al modelo de formación de opiniones. En esta ocasión, se ha modificado el valor mínimo de ϵ para que alcance valores comprendidos entre el -1 y el 1.

Usando la misma metodología que en la anterior sección, mostramos en la figura 3.14, en una media de 25 generaciones distintas, el comportamiento de la opinión media frente al valor de λ y la evolución de las varianzas de las opiniones de los individuos de la red a medida que el valor de λ avanza desde 0 hasta 2, permitiéndose de nuevo que ϵ alcance valores de hasta -1. Percibimos que, de nuevo, los resultados obtenidos, si bien presentan una forma similar a la presentada en la figura 3.6, contienen

algunas diferencias, observándose que el comportamiento de la opinión media es mucho más suave que en el caso en el que el valor mínimo de ϵ es 0. Notamos que, con estas condiciones, el valor en el que se produce el punto de ruptura de simetría, a pesar de no ser tan pronunciado como en el caso de la sección 3.2, podemos marcarlo en $\lambda_c = 0,88$, al tratarse del punto en el que se produce un mayor aumento en el valor de la opinión media.

Además puede notarse que el rango de valores para λ para los cuales la opinión se encuentra ampliamente distribuida dentro del espectro contemplado es mucho mayor que en el caso anterior, observándose altos valores para la varianza de las opiniones para los casos $\lambda \in [0,88, 1,1]$.

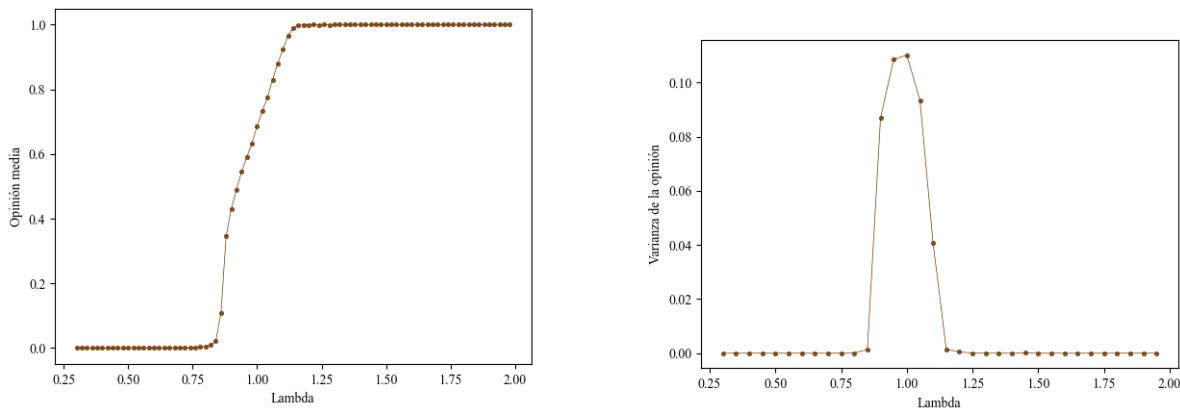


Figura 3.14: Comportamiento de la evolución media y de la varianza de la opinión respecto del valor de λ . En esta ocasión, se ha modificado el valor mínimo de ϵ para que alcance valores comprendidos entre el -1 y el 1.

EL PAPEL DE LOS MEDIOS DE COMUNICACIÓN EN LA FORMACIÓN DE OPINIONES

Hemos analizado hasta ahora los resultados obtenidos al incorporar el modelo cinético de gases ideales a una red social con sus diversas consideraciones. En este modelo se ha supuesto que todos los nodos, que se corresponden con componentes de la red social, tienen las mismas características. Pero esto, evidentemente, no se corresponde con la estructura real de las redes sociales.

Para que dicha estructura sea más fidedigna, introducimos en nuestro modelo los medios de comunicación. Un medio de comunicación se diferencia del resto de nodos en la evolución de su opinión, ya que esta no varía independientemente de las relaciones que establezca dentro de la estructura de la red social. Esto se debe a que cada medio de comunicación dispone de una línea editorial y, por lo tanto, su opinión no variará hasta que el equipo directivo del propio medio así lo decida. Para emular el papel desempeñado por los medios de comunicación, diseñaremos una serie de nodos que, en nuestras simulaciones, no actualizarán su opinión cuando se realice la interacción con sus nodos vecinos, pero que sí influenciarán en la formación de las opiniones de estos.

4.1. Diseño de algoritmo para conexión a medios de comunicación

La pregunta que surge para desarrollar este modelo es la siguiente *¿Cómo decidimos qué medios de comunicación elige cada nodo?*

Existen distintos tipos de aproximación a este problema. La solución que proponemos está basada en el concepto de sesgo de confirmación. Este sesgo cognitivo, ampliamente estudiado en el campo de la psicología, es la tendencia de los seres humanos a dar mayor peso a aquellas informaciones que favorezcan la hipótesis propia, lo que determina que en la búsqueda de fuentes para contrastar cierta información, se dé menos importancia a aquellas que se alejen de la percepción inicial que se posea [14] [16] [23]. Para ajustarnos al comportamiento derivado del sesgo de confirmación humano, hemos decidido conferir una mayor probabilidad a que un nodo preste atención a medios de comuni-

cación con una opinión más próxima a la propia, y por lo tanto establezca una conexión. De esta forma logramos un modelo más fiable que los que presentan autores como Brooks [4] o Sirbu [21] a final de sus artículos.

El algoritmo propuesto es el siguiente:

- 1.– Creamos un grafo siguiendo el procedimiento propuesto por Leskovec de igual forma que realizábamos en las secciones previas.
- 2.– Generamos un cierto número de medios de comunicación, en este caso 10, cuyas opiniones estarán distribuidas de forma uniforme entre los valores de -1 y 1.
- 3.– Determinamos el número de periódicos a los que cada nodo presta atención, en este caso 4.
- 4.– Usaremos una función de distribución beta (ver sección 2.3) con parámetros $\alpha = 5,82$, $\beta = 5,82$. La justificación de estos valores será detallada en la sección 4.3. Esto nos proporciona una función centrada en el valor $x = 0,5$, en la que los valores más altos de la función se encontrarán en los valores más cercanos a x , tiendiendo la función a 0 a medida que nos alejemos del centro. Por cada nodo generamos tantos puntos como medios de comunicación a los que se debe conectar cada nodo hayamos determinado siguiendo la distribución beta.
- 5.– No obstante, buscamos una distribución que otorgue más probabilidades a la generación de valores cercanos a la opinión del nodo propio. Puesto que la función beta está determinada en el intervalo (0, 1), multiplicamos a cada uno de los números generados por 4, le restamos dos unidades y le sumamos el valor de la opinión del nodo en cuestión. Esta transformación está mostrada con dos ejemplos en la figura 4.1.

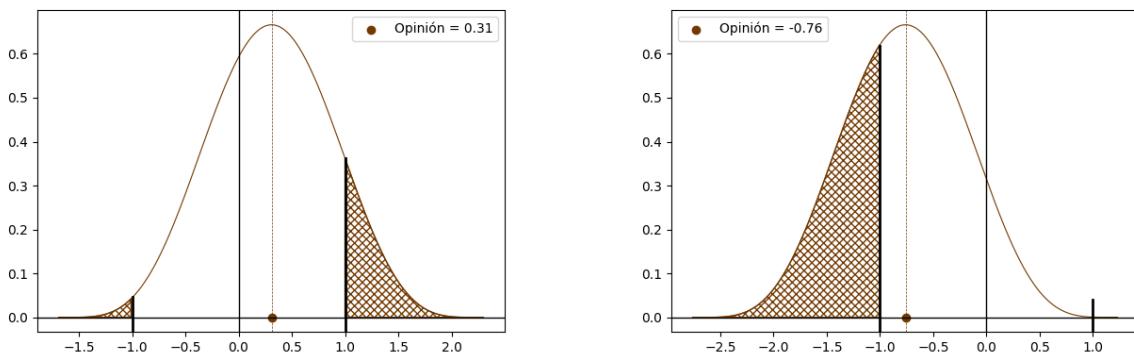


Figura 4.1: Función beta tras realizar la transformación explicada y ser centrada en la opinión del nodo pertinente. En este caso, como ejemplo, hemos establecido que la opinión del nodo sea 0,31 y -0,76 respectivamente. En caso de que el número generado cayera dentro de las áreas sombreadas, al estar fuera del rango de opiniones que contemplamos, sería necesario repetir su generación.

- 6.– En caso de que los puntos generados caigan en valores fuera de nuestro rango de opiniones (en nuestro caso esto se produce cuando sean menores que -1 o mayores que 1), volvemos a generar el punto hasta que se consigan valores dentro del rango contemplado.
- 7.– Por cada uno de esos puntos distribuidos en el rango de opiniones calculamos las distancias a las opiniones de cada uno de los medios de comunicación del sistema y establecemos una arista entre el medio de comunicación cuya opinión presente la menor distancia a la opinión del nodo y el propio nodo,

input : Número de medios a los que se conecta cada individuo, número de medios totales, una lista con los medios
output: Un grafo con cada nodo conectado a un numero determinado de medios de comunicación en función del valor de su opinión

```

1 #NUMMEDIOS=10;
2 #NUMMEDIOSCONEX=4;
3 G = grafoLeskovec ();
4 arrayPrensa = [];
5 for i in range ( NUMMEDIOS): do
6 | arrayPrensa.append( nodoPrensa)
7 end
8 for nodo in G: do
9 | if nodo not Prensa: then
10 | | numerosGenerados = [];
11 | | for i in range ( NUMMEDIOSCONEX): do
12 | | | AUX = True;
13 | | | while aux: do
14 | | | | #Generamos un número que siga una distribución beta 1, 1;
15 | | | | numBeta = beta (1,1);
16 | | | | #Realizamos la transformación referida en el punto 5;
17 | | | | numBeta = numBeta *4-2+ nodo.opinion;
18 | | | | #Comprobamos si el número generado está dentro del rango de opiniones. Si no, repetimos el
19 | | | | proceso;
20 | | | | if numBeta >-1 and numBeta <1 then
21 | | | | | AUX = False;
22 | | | end
23 | | | numerosGenerados.append( numBeta);
24 end
25 #Calculamos el medio de comunicación más cercano a cada uno de los puntos generados;
26 for num in numerosGenerados: do
27 | distancias = [];
28 | for medio in listaMedios: do
29 | | distancias.append( abs ( medio.opinion - num));
30 | end
31 | G.addEdge( nodo, listaMedios[ indexOf( min ( distancias))]);
32 | end
33 end
34 end

```

Algoritmo 4.1: En este algoritmo, que usa como base del mismo la función de distribución beta, conseguimos una nueva forma para que cada individuo establezca aristas con medios de comunicación de forma proporcional a la semejanza de la opinión de estos y la del propio individuo.

4.2. Precisión del modelo

Puede surgir la duda razonable sobre la precisión del sistema propuesto. Para tratar de responder a esta pregunta, vamos a usar una idea similar a la del estimador del error cuadrático medio (*ECM*), que aporta una visión general del error producido, de media, con el modelo propuesto.

Este estimador suele ser utilizado para determinar el valor del error medio a la hora de comparar una serie de mediciones con la media obtenida o el valor esperado. Sin embargo, nosotros no vamos a considerar esta técnica para hallar un error como tal, pues no disponemos de una base muestral para cotejar los resultados, sino que emplearemos la misma fórmula que el estimador, comprobando la distancia entre las opiniones de los medios de comunicación a los que cada nodo se encuentra conectado y la opinión del propio nodo, de aquí en adelante llamaremos a este valor, definido en la ecuación 4.1 como *DCM*, o Distancia Media Cuadrática.

Por ello, y suponiendo que el número de medios de medios de comunicación al que cada individuo de nuestra red social presta atención se mantiene siempre constante, $m = 4$, obtenemos la figura 4.2, en la que se presenta la *DCM* a la hora de elegir los medios de comunicación a los que conectarse en función del número de medios disponibles en la red. Formalmente la distancia cuadrática media se calcula de la siguiente forma:

$$DCM = \frac{\sum_{j \in J} \frac{\sum_{i \in I_j} (O_j - O_i)^2}{|I_j|}}{|J|}, \quad (4.1)$$

donde J representa el conjunto de nodos de nuestro grafo, I_j el conjunto de medios de comunicación a los que el nodo j-ésimo está conectado, O_i la opinión de cada uno de esos medios y O_j la opinión del propio nodo. Tomaremos como *DCM* el valor medio obtenido tras efectuar un total de 15 repeticiones.

Como se puede ver en la figura 4.2, tras unos resultados iniciales en los que el valor medio del error fluctúa ligeramente, este se estabiliza en torno al valor $DCM = 0,24$ cuando cada individuo puede elegir entre un número amplio de medios de comunicación.

No obstante, no se puede asegurar con total seguridad que el valor $m = 4$ sea el que más se asemeje con el comportamiento de los usuarios de las redes sociales en la vida real, por ello realizaremos este cálculo y tomaremos el valor medio de *DCM* para cada valor de m .

Tras realizar las pertinentes simulaciones, podemos observar un comportamiento similar para diferentes valores. Todos los casos presentan valores de *DCM* con una moderada dispersión cuando el número de medios de comunicación disponibles es bajo. Sin embargo a medida que este número aumenta, todas las mediciones acaban concentrándose en valores cercanos a $DCM = 0,24$. En la tabla 4.1 se presentan los diferentes resultados de la *DCM* obtenida en función del número de medios

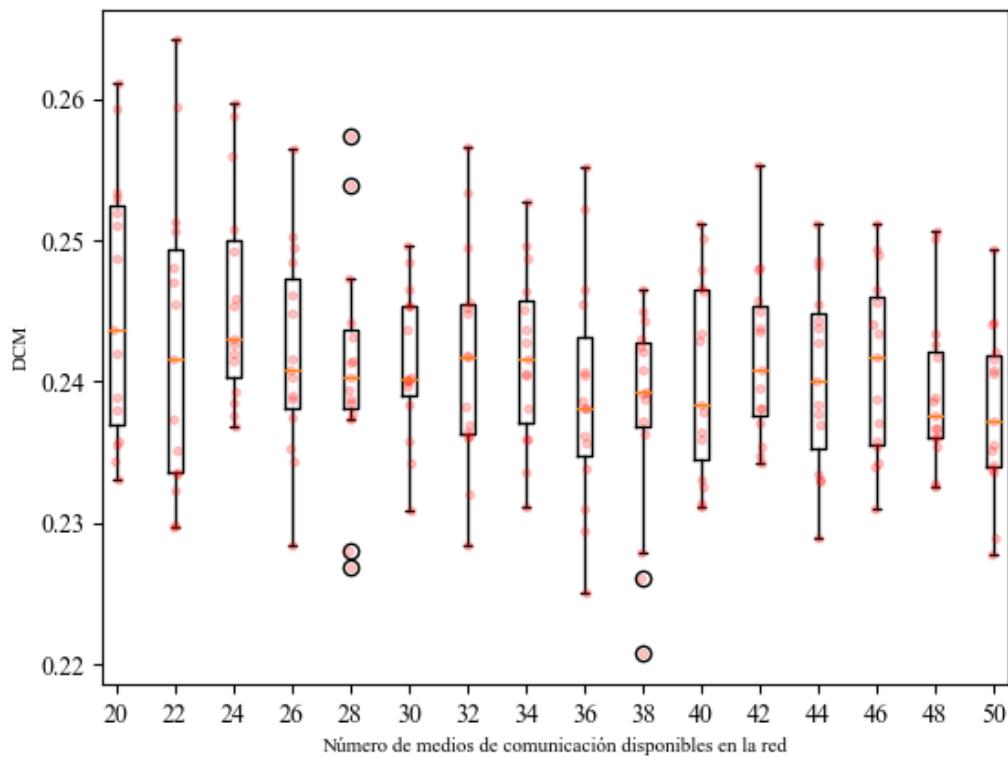


Figura 4.2: Distancia cuadrática media cuando cada individuo se conecta a 4 medios de comunicación. Puede observarse el valor de los distintos cuartiles, así como del valor máximo y mínimo, pudiéndose determinar que se trata de una distribución centrada, por lo que la *DCM* puede ser utilizado para determinar la precisión del modelo.

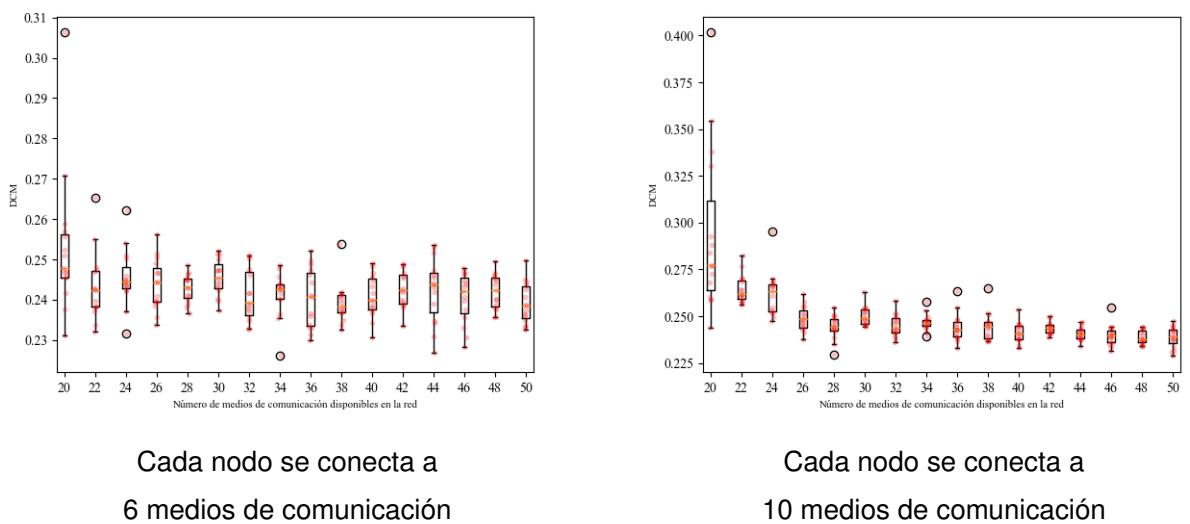


Figura 4.3: Distancia cuadrática media cuando cada individuo se conecta a 6 y 10 medios de comunicación respectivamente.

de comunicación existentes en el sistema para cada valor de m .

		Número de medios de comunicación totales					
		20	24	30	34	40	50
Valor de m	4	0.243	0.243	0.240	0.241	0.238	0.237
	6	0.248	0.244	0.245	0.242	0.24	0.238
	8	0.256	0.249	0.243	0.241	0.242	0.241
	10	0.276	0.262	0.248	0.245	0.24	0.238
	12	0.301	0.275	0.249	0.246	0.242	0.238
	14	0.34	0.293	0.257	0.248	0.241	0.241

Tabla 4.1: En la tabla se observa el valor medio de distancia cuadrática media, para cada una de las combinaciones entre medios de comunicación totales y medios de comunicación a los que se conecta cada nodo.

Destaca especialmente la última columna de la tabla 4.1, pues nos permite afirmar que cuando el número de medios disponibles, para que cada individuo les preste atención es elevado, el valor de la *DCM* es independiente del número de medios a los que está conectado. Además, vemos que, salvo muy ligeras fluctuaciones en los casos en los que el número de medios de comunicación totales es muy cercano al valor de m , todas las combinaciones posibles alcanzan valores similares de *DCM*.

4.3. Valor de los parámetros de la función beta

En la descripción del algoritmo propuesto en la sección 4.1, en concreto en el paso 4, hemos visto que el valor de los parámetros $\alpha = \beta = 5,82$. Este valor no ha sido considerado de forma aleatoria, sino que se corresponde con los resultados arrojados de un estudio que detallamos a continuación.

Cómo hemos visto en la sección 2.3, la función beta adopta diferentes gráficas en función de los parámetros α y β . Estamos interesados en que nuestra función de distribución este centrada en la opinión de cada nodo, pues debe tener igual de probabilidades de conectarse a un medio de comunicación cuya opinión que se encuentre a la misma distancia, independientemente de si esta se halle a la izquierda o a la del propio nodo. Por ello, y por las características de la función beta, los parámetros de esta deben tener el mismo valor: $\alpha = \beta = \gamma$. Como hemos visto en la figura 2.6, la forma de la función, a medida que el valor de γ aumenta, varía, observándose mayores picos en las zonas cercanas al punto medio. Pero ¿Cómo podemos determinar el parámetro γ para que el algoritmo propuesto se ajuste en mayor medida a lo observado en la vida real?

Para responder a esta pregunta vamos a usar los datos recabados en la encuesta lanzada (Para más información consultar Apéndice A). Nos va a interesar en concreto el valor medio de dispersión de la opinión de los medios elegidos respecto de la opinión propia. Vemos que este valor en la encuesta es de $DCM \approx 0,24$. Este resultado coincide con el valor al cuál tienden los valores de *DCM* a medida

que aumenta el numero de medios de comunicación disponibles para ser elegidos en la red social, tal y como puede verse en la última columna de la tabla 4.1.

Con esto, y tras realizar una serie de simulaciones con el algoritmo propuesto para hallar el valor medio de DCM para cada uno de los parámetros, buscamos este el valor del eje de las coordenadas. Estos resultados pueden observarse en la figura 4.4. Vemos que, numéricamente, este se corresponde con $\gamma \approx 5,82$. Determinamos así, el valor de nuestros parámetros para generar el algoritmo.

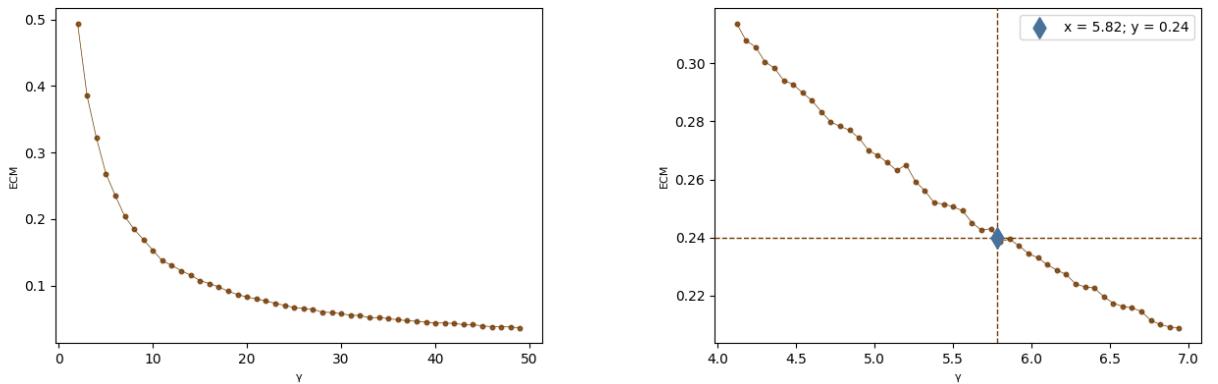


Figura 4.4: DCM del algoritmo propuesto en función del valor de γ . Valores tomados con la media de un total de 25 repeticiones con 600 nodo, generándose para cada uno 4 puntos siguiendo el algoritmo. En la subfigura (b) se observa el detalle de la subfigura (a) en valores en torno a $x = 5,82$.

4.4. Resultados obtenidos

Tras la explicación, y la pertinente justificación, de este algoritmo, procedemos a realizar una serie de simulaciones tal y como hicimos en el punto 3.2, con redes sociales compuestas por un total de 600 individuos y 15 iteraciones.

Los resultados mostrados en la figura 4.5 ponen de manifiesto ciertas diferencias con los de la figura 3.5. Se deben a la existencia de un mayor numero de nodos, que representan posiciones intermedias, es decir, no extremas, cualquiera que sea el valor de λ (apego a la opinión propia) considerado. Estos nodos corresponden a medios de comunicación, cuyas opiniones son invariantes, y por lo tanto, independientes del valor de λ . Concretamente a los medios que tienen una opinión intermedia. Se prevé que estos medios hagan de atractor hacia su opinión, impidiendo que las opiniones se sitúen en los extremos del espectro contemplado de una forma tan pronunciada.

Siguiendo la metodología del capítulo 3, representamos en la figura 4.6 la varianza de la opinión así como el valor de la opinión media respecto de λ . Para obtener la gráfica se ha tomado el valor promedio de un total de 25 repeticiones, cada una con grafos distintos para que los resultados obtenidos no dependan de la topología generada. Es en estos resultados donde encontramos las principales

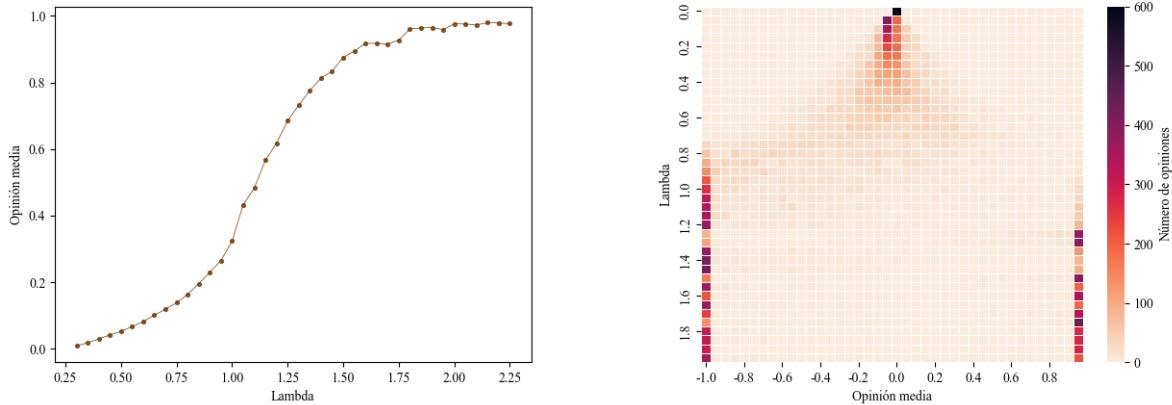


Figura 4.5: Comportamiento de la opinión media y de la distribución de las opiniones al añadir a la estructura de un grafo de Leskovec estudiada en el capítulo 3 una serie de individuos que cumplen el papel de los medios de comunicación. Se observan diferencias sustanciales en la forma de ambas gráficas, observándose un comportamiento más heterogéneo de las partículas del sistema.

diferencias con respecto a los mostrados en el punto 3.2. En primer lugar, los resultados muestran, en efecto un cambio notable en el avance de la opinión media a medida que avanza el valor de λ , no produciéndose ni el consenso ni el posicionamiento de las opiniones en los extremos visto en figura 3.6 para ningún valor de λ . Sí que se produce una transición desde estados en los que la opinión media es cercana a 0 hacia unos cercanos a 1, pero esta vez no se alcanza ninguno de los valores. Además esta transición se produce de una forma mucho más suave, siendo imposible determinar un punto de ruptura de simetría λ_c .

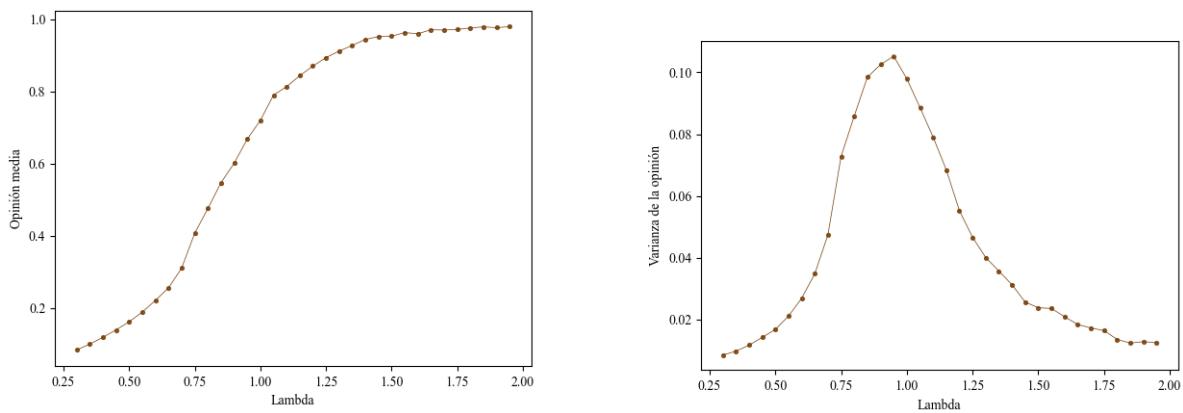


Figura 4.6: Comportamiento de la opinión media y de la varianza de la opinión respecto del valor de λ al añadir individuos a nuestra red que cumplen el papel de los medios de comunicación.

Para analizar los motivos de este fenómeno, observamos en la figura 4.6 otra importante diferencia con la figura 3.6. Es claro que ahora la distribución de opiniones es muy variada, encontrándose

distribuidas de manera más dispersa en valores cercanos a $\lambda \sim 1,02$. Esto contrasta radicalmente con los resultados obtenidos sin la presencia de los medios de comunicación en nuestra red social, pues en estos casos, todas las opiniones se hallaban completamente concentradas salvo para un rango de valores muy pequeño en torno a $\lambda = 0,68$.

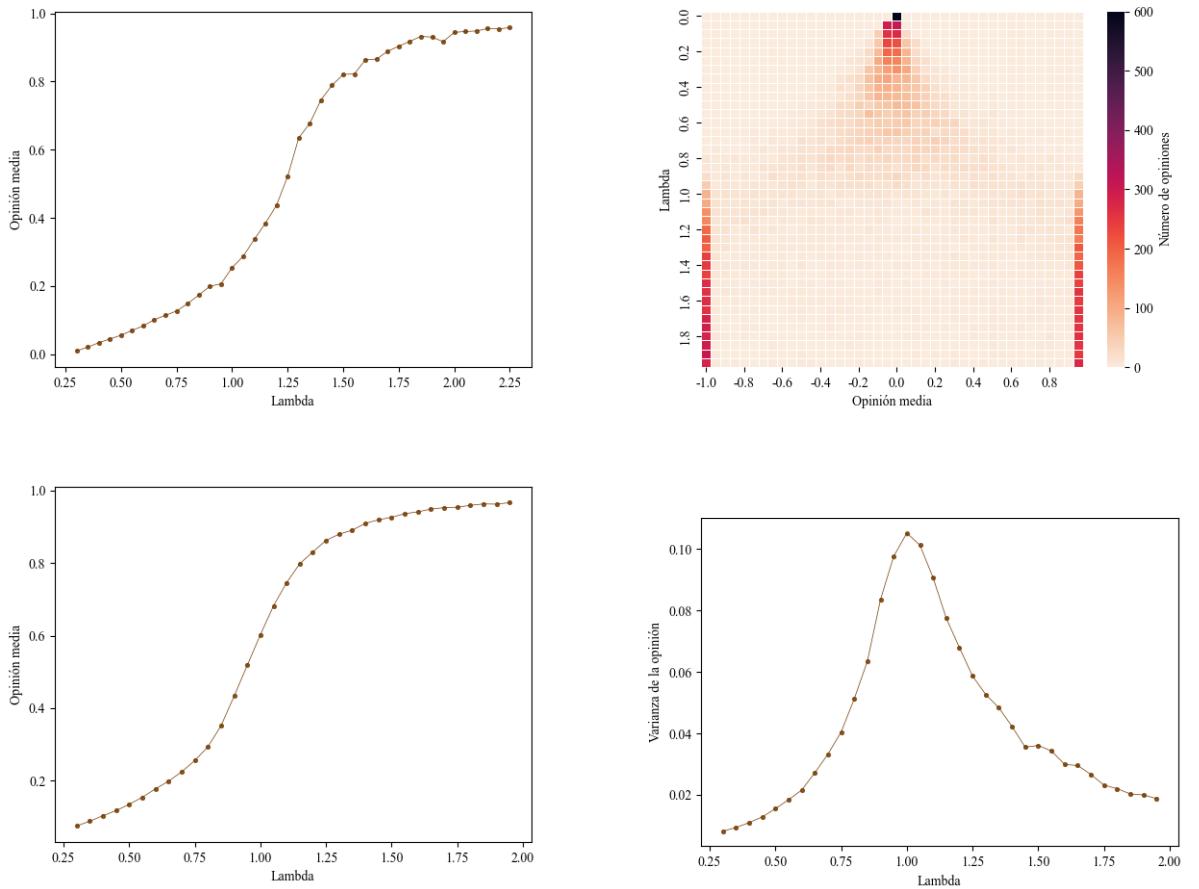


Figura 4.7: En la línea superior se presenta el mapa de calor de la variación de la opinión a lo largo de los diferentes valores de λ y su respectivo gráfico de opinión media. En la línea inferior presentamos, tras un total de 25 repeticiones, el valor de la opinión media respecto al valor de λ y varianza de la opinión.

Este es un resultado notable ya que indica que, gracias a la presencia de los medios de comunicación en una red social, el consenso total no se alcanzará prácticamente nunca, sin embargo, también será mucho más complicado que el sistema llegue a estados en los que la opinión de los componentes de la red se ubique completamente en posiciones extremas (de hecho para $\lambda = 2$, el valor que alcanza el valor absoluto de la opinión media es 0,97). Por ello, podemos afirmar que, los medios de comunicación, sirven de elementos que posicionan las opiniones de los individuos, que conforman una red social, en estados intermedios entre el consenso y la polarización. Y además, lo hacen de una forma mucho más dispersa en todo el rango de opiniones.

En la sección 3.4 vimos que el valor mínimo del parámetro ϵ en la ecuación 2.1 era un elemen-

to que debía tenerse en cuenta, pues en función del mismo se obtenían unos resultados diferentes del comportamiento general del sistema. Procedemos a realizar las mismas simulaciones modificando únicamente este parámetro, permitiendo que adquiera valores pertenecientes al intervalo $[-1,1]$, mostrando el resultado en la figura 4.7.

Nuevamente, se presentan una serie de curvas de evolución hacia opiniones extremas mucho menos pronunciadas que las vistas en la figura 3.5. La posibilidad de la adopción de valores negativos para la variable ϵ ha suavizado aún más el comportamiento de la opinión media respecto del valor de λ . De hecho, y tomando como referencia la figura 4.6, se puede observar un comportamiento de la varianza de la opinión en el que se alcanzan cotas mayores que en cualquiera de los casos anteriores. Nuevamente el punto en el que las opiniones de los individuos se encuentran más distribuidas se corresponden con valores de λ próximos a 1.

Llama también la atención de manera especial el hecho de que, en este caso, la posibilidad de que el parámetro ϵ alcance valores negativos, haya provocado que el valor máximo de la opinión media, si bien tiende a 1, nunca alcance dicho valor, potenciando aún más el fenómeno, que ya se producía con la incorporación de los medios de comunicación en la red diseñada, al producirse una mayor dispersión en las opiniones de la población. De hecho, fijándonos en el gráfico superior derecho de la figura 4.7, pese a que nuestra red está compuesta por 600 individuos, en los casos en los que el valor de λ es alto, con las opiniones que se hallan en los extremos no suman la totalidad de individuos de nuestra red, estando el restante distribuidos entre las opiniones intermedias.

CONCLUSIONES Y LIMITACIONES

5.1. Conclusiones

A lo largo del desarrollo de esta memoria se han analizado distintas alternativas para mejorar el trabajo descrito en el artículo de Mehdi Lallouache [12]. Hemos empleado diferentes estrategias de formación de redes sociales para estudiar la evolución de las opiniones dentro de una red social a medida que se producen sucesivas interacciones, tal y como se indicaba en el primero de los objetivos de este trabajo. Todo este trabajo ha sido modelado por una ecuación que permite el intercambio de opiniones de forma simultánea entre dos nodos que son adyacentes en el grafo que representa a la red social. Tras estudiar los valores de λ para los cuales el sistema transita desde un estado de consenso general en el que la opinión de todos los individuos de la red converge en el 0 a uno en el que todas, o prácticamente todas, las opiniones se sitúan en posiciones extremas. Este cambio, de forma similar al estudiado en [12] ($\lambda_c = 0,68$), se produce de forma repentina en $\lambda_c \sim 0,7$.

Hemos visto también que usando otras estrategias de creación de grafos, como Erdős-Rényi o Barabási, los resultados obtenidos han sido prácticamente iguales: $\lambda_c = 0,67$ y $\lambda_c = 0,68$. En último lugar se han analizado una serie de consideraciones adicionales, tales como ampliar el valor mínimo del parámetro ϵ de la ecuación 2.1, estudiar el valor de λ_c en función de la conectividad media de cada uno de los nodos del grafo, la tolerancia o el número de individuos de la red social, obteniéndose como resultados que para todos los casos el valor de λ_c hallado siempre es similar a la encontrada en secciones anteriores, excepto para las simulaciones en las cuales variamos el valor mínimo del parámetro ϵ .

Nos resulta relevante mencionar las variaciones que sufren los resultados obtenidos tras realizar las pertinentes simulaciones al variar el valor mínimo del parámetro ϵ . Esta variable, que es generada distinta para cada par de individuos en cada transferencia de opinión, representa el poder de influencia que tiene un agente dado en un momento determinado sobre la opinión de un determinado individuo. Al permitir que esta variable adopte parámetros negativos, estamos representando que pueden darse situaciones en las que una persona nos aleje de su opinión debido a que la influencia que tiene sobre nuestra opinión en ese momento es negativa. Lo más relevante que la inclusión de este factor provoca,

es la eliminación de un punto de ruptura de simetría y el paso más suave y lento desde un sistema en consenso con opiniones en torno al 0 hacia un sistema en el que las opiniones se concentran en los valores extremos del rango contemplado.

También hemos introducido, siguiendo el segundo de los objetivos propuestos para este trabajo, una nueva variante al modelo descrito, la incorporación de individuos cuyas opiniones no varían a lo largo del tiempo, los medios de comunicación. Tras proponer un modelo de conexión de cada individuo a un número determinado de medios de comunicación, basado en una serie de resultados obtenidos entre varios encuestados, se ha observado que la conclusión obtenida en la sección anterior varía ligeramente, no pudiéndose determinar de una forma tan clara un valor para λ_c , ya que la transición entre estados se produce de una forma mucho más paulatina. Con esta variante no se observa en ningún momento que la opinión media de la red sea 0 o 1, al contrario que en los casos en los que no existía el papel de los medios de comunicación. Este hecho se debe a que se puede observar una mayor dispersión en la distribución de las opiniones de los componentes de la red social.

Con esto podemos afirmar que el papel de los medios de comunicación dentro de una red social es el de disminuir la formación de grandes clústeres de nodos que tengan la misma opinión, distribuyéndose estas de una forma mucho más dispersa, lo que provoca que la red no llegue a un consenso para valores de λ bajos, pero que tampoco llegue a posiciones extremas para valores altos de λ .

5.2. Limitaciones y trabajo futuro

Como posibles líneas de trabajo futuro, se propone realizar un estudio de la formación de opiniones de la red social en presencia de medios de comunicación en el que el valor de la variable λ no sea el mismo para todos los individuos, teniendo eso sí, el de cada individuo un valor fijo a lo largo de la propia simulación, pues al final se trata de un parámetro que determina el apego a la opinión propia, y no tiene por qué ser igual en todos los sujetos.

Siguiendo también la idea de [20], se propone realizar un estudio que represente el papel de la influencia en las redes sociales. En él, el valor de la influencia, entre sus vecinos, de los nodos con mayor grado de conectividad, será mayor que el de los nodos con pocos vecinos. Se trataría de representar la figura de los *influencers* dentro de las redes sociales.

En último lugar, constantemente a lo largo del desarrollo de esta memoria hemos estado trabajando con grafos estáticos. Sin embargo, esto no se corresponde con el desarrollo que tienen las redes sociales en la vida real. Por ello, se propone realizar un estudio similar al presentado, pero con cambios en la topología del grafo que modela la red, de forma dinámica.

BIBLIOGRAFÍA

- [1] BARABASI, A.-L., AND ALBERT, R. Emergence of scaling in random networks. *Science* 286, 5439 (oct 1999), 509–512.
- [2] BAYES, T. An essay towards solving a problem in the doctrine of chances. *Phil. Trans. of the Royal Soc. of London* 53 (1763), 370–418.
- [3] BERNOULLI, D. *Hydrodynamica, sive de viribus et motibus fluidorum commentarii*. sumptibus Johannis Reinholdi Dulseckeri, 1738.
- [4] BROOKS, H. Z., AND PORTER, M. A. A model for the influence of media on the ideology of content in online social networks. *Phys. Rev. Research* 2 (Apr 2020), 023041.
- [5] BRUGNA, C., AND TOSCANI, G. Kinetic models of opinion formation in the presence of personal conviction. *Physical Review E* 92, 5 (nov 2015).
- [6] CASTELLANO, C., MUÑOZ, M. A., AND PASTOR-SATORRAS, R. Nonlinear q-voter model. *Physical Review E* 80, 4 (oct 2009).
- [7] CHAKRABORTI, A., AND CHAKRABARTI, B. Statistical mechanics of money: how saving propensity affects its distribution. *The European Physical Journal B* 17, 1 (aug 2000), 167–170.
- [8] ERDÖS, P., AND RÉNYI, A. On random graphs i. *Publicationes Mathematicae Debrecen* 6 (1959), 290.
- [9] FERNÁNDEZ-GRACIA, J., SUCHEKI, K., RAMASCO, J. J., MIGUEL, M. S., AND EGUILUZ, V. M. Is the voter model a model for voters? *Physical Review Letters* 112, 15 (apr 2014).
- [10] HORNE, B. D., AND ADALI, S. This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news, 2017.
- [11] ISING, E. Contribution to the Theory of Ferromagnetism. *Z. Phys.* 31 (1925), 253–258.
- [12] LALLOUACHE, M., CHAKRABORTI, A. S., CHAKRABORTI, A., AND CHAKRABARTI, B. K. Opinion formation in kinetic exchange models: Spontaneous symmetry-breaking transition. *Physical Review E* 82, 5 (nov 2010).
- [13] LESKOVEC, J., BACKSTROM, L., KUMAR, R., AND TOMKINS, A. Microscopic evolution of social networks. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (New York, NY, USA, 2008), ACM, pp. 462–470.
- [14] LORD, C., ROSS, L., AND LEPPER, M. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology* 37 (11 1979), 2098–2109.
- [15] MEDIA, A. F., 2021.
- [16] MODGIL, S., SINGH, R., GUPTA, S., AND DENNEHY, D. A confirmation bias view on social media induced polarisation during covid-19. *Information Systems Frontiers* (11 2021).
- [17] MUÑOZ, M. A., 2019.

- [18] R., H., AND U., K. Opinion dynamics and bounded confidence. models, analysis and simulation. *Journal of Artificial Societies and Social Simulation* 5, 3 (2002), 1–33. (Descargar).
- [19] REDNER, S. Reality-inspired voter models: A mini-review. *Comptes Rendus Physique* 20, 4 (may 2019), 275–292.
- [20] SEN, P. Phase transitions in a two-parameter model of opinion dynamics with random kinetic exchanges. *Physical Review E* 83, 1 (jan 2011).
- [21] SÎRBU, A., PEDRESCHEI, D., GIANNOTTI, F., AND KERTÉSZ, J. Algorithmic bias amplifies opinion fragmentation and polarization: A bounded confidence model. *PLOS ONE* 14, 3 (03 2019), 1–20.
- [22] TOSCANI, G. Kinetic models of opinion formation, 2006.
- [23] WANG, Q., AND JEON, H. J. Bias in bias recognition: People view others but not themselves as biased by preexisting beliefs and social stigmas. *PLOS ONE* 15, 10 (10 2020), 1–18.
- [24] YULE, G. U., AND FILON, L. N. G. Karl pearson. 1857-1936. *Obituary Notices of Fellows of the Royal Society* 2, 5 (1936), 73–110.

APÉNDICES

RESULTADOS DE LA ENCUESTA REALIZADA

Presentamos en este apéndice el estudio llevado a cabo para complementar los resultados presentados en el capítulo 4. Esta encuesta ha sido realizada persiguiendo el objetivo de estudiar la dispersión en la selección de medios de comunicación que cada usuario toma como referencia. Puesto que resulta una complicación innecesaria realizar un estudio basándonos en la posición de cada usuario y medio de comunicación con respecto a varios asuntos, hemos decidido aprovechar el trabajo realizado por Ad Fontes Media y fijarnos únicamente en la posición política que determina la línea editorial de cada uno de los medios de comunicación españoles. De esta forma podemos desarrollar un análisis monodimensional de los parámetros de dispersión a la hora de seleccionar a qué medios de comunicación se conecta cada usuario teniendo en cuenta la opinión de los mismos y del usuario en cuestión.

A.1. El estudio de Ad Fontes Media

El primer paso en nuestro experimento es el de posicionar cada uno de los medios de comunicación existentes en España en el espectro de opinión (en este caso político), entre -1 y 1. Para llevar a cabo este paso usaremos como referencia el estudio de Ad Fontes Media [15], una empresa estadounidense de periodismo de vigilancia. Han cobrado renombre dentro de la comunidad por el desarrollo de una serie de gráficos en los cuales posicionan en dos ejes cada uno de los medios de comunicación existentes en función de su posicionamiento político y la credibilidad de sus noticias.

Esta empresa lanzó en 2021 el primer estudio en el que se analizaba el papel de los 30 principales medios de comunicación escrita en España, situando a cada uno de estos como puede verse en la figura A.1 en un gráfico de dos coordenadas, atendiendo a la fiabilidad general del medio de comunicación y a su posicionamiento político de manera general.

Para este estudio analizan un total de 450 artículos de 30 distintos medios, teniendo en cuenta facetas de estos como le lenguaje utilizado dentro del artículo, sobre qué objetivos focaliza su incidencia el artículo o la objetividad general del artículo.

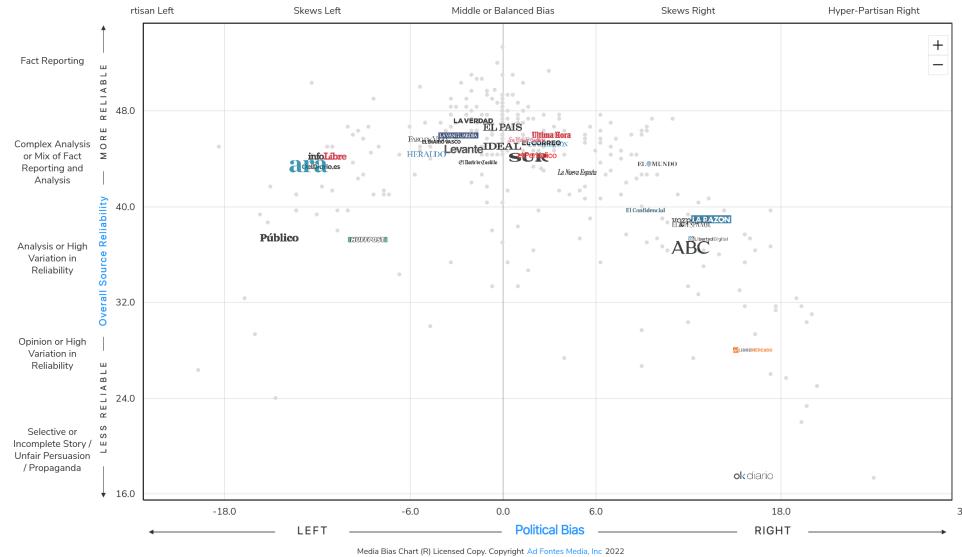


Figura A.1: Gráfico de la distribución de los medios de comunicación escrita españoles. El eje X se corresponde con el posicionamiento ideológico del medio mientras que el Y con la credibilidad que se le asigna a dicho medio. Cada punto se corresponde con uno de los 450 artículos analizados en el estudio de Ad Fontes Media. Figura obtenida de [15].

A.2. Posicionamiento de los medios de comunicación

Con el fin de complementar el algoritmo desarrollado en el punto 4.1, vamos a interesarnos únicamente en el posicionamiento político de cada medio de comunicación español. Esto se corresponde con el eje de abscisas en el gráfico de Ad Fontes Media. Para transformar el rango planteado por el estudio en el que nosotros estamos interesados, $[-1, 1]$, hemos tomado como mínimo y máximos absolutas las noticias más extremas encontradas dentro del estudio, correspondiéndose estas con sendas noticias de Público¹ y Okdiario². De esta forma podemos asignar a cada uno de los medios de comunicación españoles el siguiente valor dentro del espectro de opiniones política tal y como se muestra en la tabla A.1.

A.3. Resultados de la encuesta

Tras esto se lanzó una encuesta realizada por medio de Google Forms en la que se preguntaba, de forma anónima, el posicionamiento político de cada encuestado en una escala del 1 al 10 y, dentro de la lista de medios de comunicación de la tabla A.1, la selección de los 4 que prefieren para informarse de manera genérica. Evidentemente este estudio no tiene en cuenta factores como la relevancia de cada medio de comunicación en el panorama periodístico español, pues, como parecía intuitivo, El

¹ <https://blogs.publico.es/otrasmiradas/47823/ayuso-y-las-residencias-de-ancianos-mentiras-y-ocultaciones-para-tapar-un-desastre/>

² <https://okdiario.com/espana/sanchez-presenta-cupula-del-ibex-ofensiva-antipodemios-hay-que-volver-bipartidismo-7055358>

Público	-0,56	La Verdad	-0,07	La Nueva España	0,19
Ara	-0,49	El Norte de Castilla	-0,06	El Confidencial	0,36
eldiario.es	-0,45	El País	0	El Mundo	0,39
infoLibre	-0,44	Ideal	0	ABC	0,47
Huffington Post	-0,34	La Voz de Galicia	0,06	El Español	0,47
Heraldo	-0,19	SUR	0,07	Voz Populi	0,48
Faro de Vigo	-0,19	El Periódico	0,09	Libertad Digital	0,51
El Diario Vasco	-0,15	El Correo	0,10	La Razón	0,52
La Vanguardia	-0,11	Información	0,11	Libre Mercado	0,63
Levante	-0,10	Última Hora	0,12	Okdiario	0,63

Tabla A.1: Distribución del posicionamiento de los medios de comunicación españoles en un rango de opiniones comprendido entre $[-1, 1]$.

País y El Mundo, los periódicos más importantes a nivel nacional, son los que han sido marcados por un mayor número de encuestados. Los resultados de la encuesta se muestra en la figura A.2 y A.3.

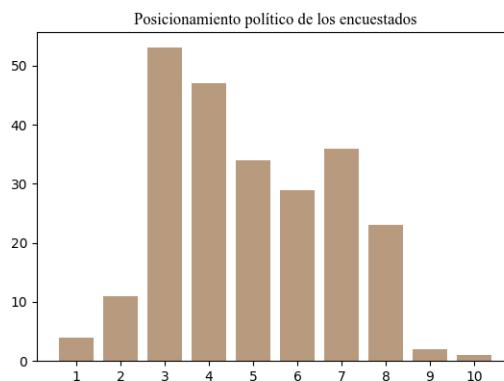


Figura A.2: Resultados obtenidos en la primera pregunta de la encuesta realizada. Se ha consultado a un total de 245 personas que han indicado su posicionamiento político, marcando 1, o 10, si consideran que siguen ideologías de extrema izquierda o de extrema derecha, respectivamente.

Con estos resultados podemos calcular, siguiendo la ecuación 4.1 la *DCM* media de los encuestado, viendo que el valor de este es $0,241482 \sim 0,24$. Veremos que este resultado nos será de gran utilidad en la sección 4.3, permitiéndonos hallar el valor parámetro γ que se ajuste con la *DCM* obtenida de forma experimental.

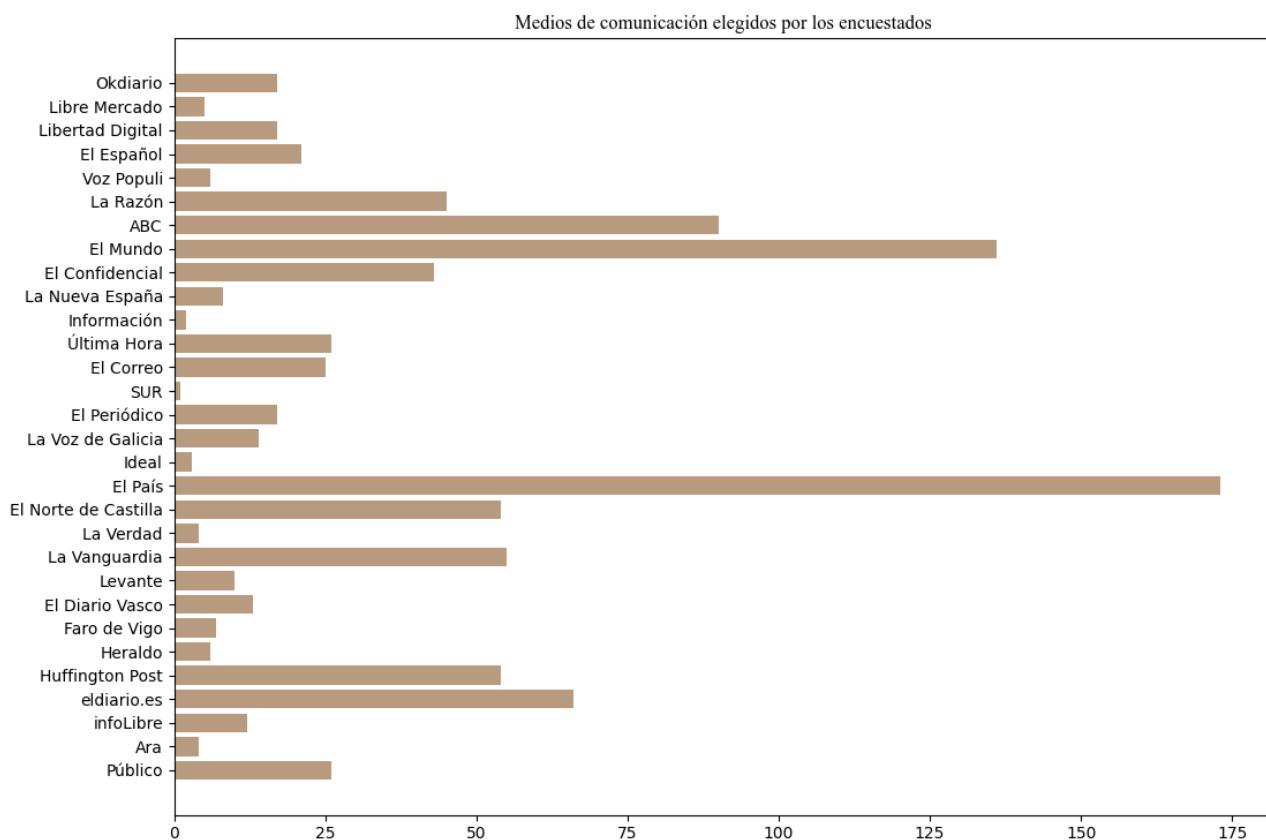


Figura A.3: Resultados obtenidos en la segunda pregunta de la encuesta realizada. Se ha consultado los 4 medios de comunicación, entre los 30 mencionados elegiría cada encuestado, mostrándose la frecuencia absoluta de cada uno de los periódicos elegidos.

