Modelización de la polarización con conjuntos difusos opuestos

Estrella García-Pelluz[1][0009-0003-4341-3956], Juan Antonio Guevara[2][0000-0003-3946-3910], Javier Montero[3][0000-0001-8333-2155], and Juan Carlos Losada[4][0000-0002-4373-603X]

[1] Faculty of Mathematics, Complutense University, 28040 Madrid, Spain
[2] Faculty of Statistical Studies, Complutense University, 28040 Madrid, Spain
[3] Faculty of Mathematics and Instituto de Matemática Interdisciplinar, Complutense University, 28040 Madrid, Spain
[4] Complex Systems Group, Universidad Politécnica de Madrid, 28040 Madrid, Spain
{estrelga, juanguev, jamonter}@ucm.es, juancarlos.losada@upm.es

Abstract. Numerosos estudios han centrado su atención en el diseño de procedimientos para detectar módulos aislados dentro de una comunidad, haciendo hincapié en el riesgo de polarización en una sociedad fracturada. En este artículo, proponemos un modelo básico para analizar la polarización afectiva, señalando la importancia de tener en cuenta tanto el apoyo (amor) como el rechazo (odio) hacia cada alternativa, como sentimientos opuestos difusos. Restringidos a dos alternativas, representamos en el espacio $[-1, +1]^2$ el odio/amor individual a cada alternativa, para entonces proponer una nueva medida de polarización social.

Keywords: Polarización · Redes Sociales · Conjuntos difusos

1 Introducción

En las sociedades democráticas actuales parece no haber espacio para largos debates, no hay oportunidad de explicar acuerdos complejos. La toma de decisiones complejas basada en múltiples criterios se está reduciendo a problemas de toma de decisiones basados en un único criterio, que suele ser binario. Un sentimiento simple no necesita argumentos sofisticados, y las ideas simples se convierten fácilmente en ideas extremas. Además, dada la rapidez con la que se propagan, resulta bastante difícil reaccionar adecuadamente (p.e, contra las fake news).

Desde el punto de vista de la comunicación, la existencia de comunidades aisladas dentro de una sociedad es un riesgo de polarización. Pero con bastante frecuencia ese aislamiento no implica polarización por sí solo. La polarización surge cuandon una parte relevante de la sociedad ama una alternativa odiando al resto de alternativas y, al mismo tiempo, otra parte relevante de la sociedad ama una alternativa diferente odiando al resto de alternativas.

El argumento central de este artículo es la necesidad de tener en cuenta dos características opuestas para detectar la radicalización individual y la polarización social ante dos alternativas, midiendo el apoyo y el rechazo hacia cada una de ellas, y que tanto el apoyo (amor) como el rechazo (odio) tienen una naturaleza difusa y graduable, lo que permite distintos grados de intensidad. Nuestro objetivo es describir la radicalización y la polarización a partir de una escala de actitud definida en el intervalo $[-1, +1]$ para cada individuo y cada alternativa.

2 Trabajos Relacionados

El elemento clave que queremos presentar en este artículo es que debemos, de alguna manera, detectar y diferenciar los mensajes positivos de los negativos, para detectar la polarización. Además, es importante reconocer que estos matices positivos y negativos admiten gradaciones. Un primer enfoque consistiría en suponer que estos mensajes están a favor o en contra de una alternativa, en cierta medida. Esta situación puede modelarse mediante una escala de actitudes definida en el intervalo $[-1, 1]$ (ver [1]). Además, debemos tener en cuenta que la polarización no surge por distribuciones específicas de conexiones positivas o negativas fuertes y débiles, sino debido a la distribución antagónica de las conexiones de apoyo y rechazo. Esta visión fue capturada en [2] donde se proponen grados de pertenencia paa modelar el grado en que un individuo puede radicalizarse hacia ambos polos ideológicos simultáneamente, pero los autores no especifican explícitamente si un grado de pertenencia nulo debe interpretarse como un rechazo explícito de ese polo o simplemente como la ausencia de radicalización hacia él. Esta distinción no es trivial. La negación de un concepto no coincide necesariamente con su opuesto. La falta de pertenencia a un polo no implica automáticamente la pertenencia al opuesto, ya que «no pertenecer» a un concepto y «pertenecer a su opuesto» representan situaciones semánticas diferentes. En nuestra opinión, la cuestión clave subyacente fue planteada originalmente por Atanassov [3], quien destacó el hecho de que un conjunto difuso no puede entenderse sin tener en cuenta sus conceptos relacionados (véanse en particular los enfoques alternativos en [4].

3 Metodología propuesta

El argumento principal de este artículo es la necesidad de considerar dos características opuestas para detectar la radicalización y la polarización, mediante la medición tanto del apoyo como del rechazo hacia cada alternativa, y que tanto el apoyo (*amor*) como el rechazo (*odio*) poseen una naturaleza difusa (*fuzzy*) graduable, permitiendo distintos grados de intensidad.

Nuestro objetivo es describir la radicalización individual y la polarización social cuando la sociedad está enfrentada a dos alternativas, mediante una escala actitudinal definida sobre el cuadrado $[-1, 1]^2$.

Los valores negativos representan odio explícito hacia la alternativa, mientras que los valores positivos representan amor explícito.

Por tanto, la opinión de un individuo i de un grupo N respecto a dos alternativas α y β puede representarse como un punto $(x_i, y_i) \in [-1, 1]^2$, siendo x_i el grado de apoyo/rechazo de cada individuo a α, e y_i el grado de apoyo/rechazo de cada individuo a β.

El objetivo de este trabajo es trasladar al contexto bipolar en el sentido de [5] al riesgo de polarización JDJ presentado en [2], considerando los grados de pertenencia a los puntos de polarización $A = (+1, -1)$ y $B = (-1, +1)$ para cada grupo.

Para cada individuo $i \in N$, los grados de pertenencia del individuo i a los puntos de polarización A y B, respectivamente, los vamos a calcular mediante una distancia normalizada en $[0, 1]$ de cada $(x, y) \in [-1, 1]^2$ a A y a B: $\mu_A(i), \mu_B(i) \in [0, 1]$.

Consideremos una función de agregación apropiadan (p.e., overlapping)

$$\varphi : [0, 1]^2 \to [0, 1]$$

y otra función de agregación (p.e., una función de grouping)

$$\psi : [0, 1]^2 \to [0, 1]$$

para combinar los grados de pertenencia en la construcción de medidas de polarización. En particular,nuestra versión bipolar del índice JDJ, denominada medida de polarización AJDJ, se define como:

$$AJDJ(Z) = \sum_{i,j \in N} \psi(\varphi(\mu_A(i), \mu_B(j)), \varphi(\mu_B(i), \mu_A(j))) \,. \tag{1}$$

donde ψ, φ son funciones de agregación definidas en el intervalo $[0, 1]$.

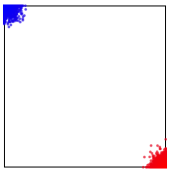
4 Resultados Preeliminares

El objetivo de esta sección es simplemente comprobar que nuestro modelo se comporrta de forma intuitiva en diferentes circunstancias, es decir, que AJDJ capta la polarización colectiva, midiendo el grado en que los individuos de la población se oponen entre sí en función de sus afiliaciones ideológicas.

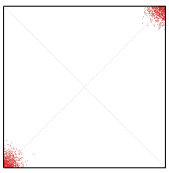
Para comprobar la intuición del índice propuesto, consideramos las siguientes funciones de overlapping y groupping (t-norma y t-conorma):

$$\varphi(a, b) = \min(a, b), \qquad \psi(a, b) = a + b - ab, \qquad a, b \in [0, 1]. \tag{2}$$

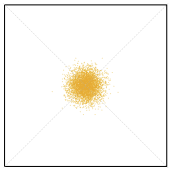
Los resultados que se obtienen para los diferentes diferentes distribuciones de opiniones individuales son:



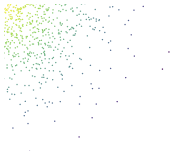
AJDJ=0.94292



AJDJ=0.23895



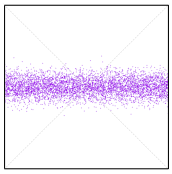
AJDJ=0.55308



AJDJ=0.14053



AJDJ=0.60041



AJDJ=0.60934

5 Conclusión

El siguiente paso inmediato, además de probar nuestra nueva medida con datos reales de redes sociales, consiste en explorar diferentes medidas de apoyo y rechazo, analizar diferentes grados de radicalización y polarización con más de dos alternativas y analizar también diferentes métodos de agregación para representar el riesgo de polarización. Nuestro modelo nos permitirá, además, medir la demanda social de una tercera alternativa.

References

1. Losada, J.C., Robles, J.M., Benito, R.M., Caballero, R.: Love and Hate During Political Campaigns in Social Networks. In: Benito, R.M. et al. (eds.) Complex Networks & Their Applications X. Studies in Computational Intelligence, vol. 1073, pp. 66–77. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-93413-2_6
2. Guevara, J.A.: Polarization Measures in Social Networks. Doctoral dissertation (in Spanish), Complutense University of Madrid (2022). `https://hdl.handle.net/20.500.14352/4170`
3. Atanassov, K.: Intuitionistic fuzzy sets. Studies in Fuzziness and Soft Computing, vol. 35 (1999). https://doi.org/10.1016/S0165-0114(86)80034-3
4. Bustince, H., Barrenechea, E., Burusco, A., Fernández, J., Rodríguez, J.T., Montero, J., Pagola, M., Gómez, D.: From Trillas' Negations and Antonyms to a Set Representation of Contradiction Within Bipolar and Other Extensions of Fuzzy Sets. Studies in Fuzziness and Soft Computing **323**, 159-177 (2015). https://doi.org/10.1007/978-3-319-18606-1_15
5. Montero, J., Bustince, H., Franco, C., Rodríguez, J.T., Gómez, D., Pagola, M., Fernández, J., Barrenechea, E.: Paired structures in knowledge representation. Knowledge-Based Systems **100**, 50-58 (2016). https://doi.org/10.1016/j.knosys.2016.02.003