Quantifying Cultural Diversity in Social Networks: A Community Embedding Approach.

Defining Diversity Measures through Graph and Machine Learning Techniques.

Abi Oppenheim[1], Federico Albanese[2,3], and Esteban Feuerstein[1,3]
{aoppenheim, falbanese, efeuerst}@dc.uba.ar

[1] Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina.
[2] Instituto de Cálculo, CONICET Universidad de Buenos Aires,Buenos Aires, Argentina.
[3] Instituto de Ciencias de la Computación, CONICET Universidad de Buenos Aires, Buenos Aires, Argentina.

Resumen The homophily phenomenon in social networks causes users to interact primarily with others who share their interests and cultural backgrounds, leading to the formation of "echo chambers" [1–3].
The notion of *cultural diversity* among users and communities becomes relevant in this context. While previous studies have investigated diversity in interaction graphs, to the best of our knowledge, none have explored the degree of diversity based on community embedding, which has been proven effective in measuring the positioning of communities in various social dimensions [4–7].
Building on the work of [7], we propose characterizing and measuring diversity through an innovative algorithm based on community embedding. We propose a novel algorithm based on community embedding to characterize and measure diversity. Our approach builds upon prior work on diversity in social media and involves iteratively updating values for the diversity of communities and individual users.
To demonstrate the effectiveness of our algorithm, we conduct a case study analyzing over over 800 million posts in 9 million discussion subreddits of different ethnic groups on Reddit. Next, we generated embeddings for each community using community2vec [8] and developed algorithms to quantify cultural diversity based on these embeddings.

Keywords: Machine learning · Social Media · Reddit · Community Embedding · Diversity.

References

1. QUATTROCIOCCHI, Walter; SCALA, Antonio; SUNSTEIN, Cass R. Echo chambers on Facebook. Available at SSRN 2795110, 2016.

2. MILLS, Judson; ARONSON, Elliot; ROBINSON, Hal. Selectivity in exposure to information. The Journal of Abnormal and Social Psychology, 1959, vol. 59, no 2, p. 250.
3. Barbera, P.; Jost, J. T.; Nagler, J.; Tucker, J. A.; and Bonneau, R. 2015. Tweeting from left to right: Is online political communication more than an echo chamber? Psychological science 26(10): 1531–1542.
4. CINELLI, Matteo, et al. The echo chamber effect on social media. Proceedings of the National Academy of Sciences, 2021, vol. 118, no 9, p. e20233.
5. QI, Guo-Jun; AGGARWAL, Charu C.; HUANG, Thomas. Community detection with edge content in social media networks. En 2012 IEEE 28th International conference on data engineering. IEEE, 2012. p. 534-545
6. HOHMANN, Marilena; DEVRIENDT, Karel; COSCIA, Michele. Quantifying ideological polarization on a network using generalized Euclidean distance. Science Advances, 2023, vol. 9, no 9, p. eabq2044.
7. WALLER, Isaac; ANDERSON, Ashton. Quantifying social organization and political polarization in online platforms. Nature, 2021, vol. 600, no 7888, p. 264-268.
8. MARTIN, Trevor. community2vec: Vector representations of online communities encode semantic relationships. En Proceedings of the Second Workshop on NLP and Computational Social Science. 2017. p. 27-31.