COMPUTATIONAL SOCIAL SCIENCE

Modeling and combating misinformation spread

Nat. Hum. Behav. https://doi.org/10.1038/s41562-022-01388-6 (2022)

The rapid spread of digital misinformation has been shown to destabilize elections and to promote erroneous practices to combat COVID-19. Due to the pervasive nature of online misinformation, a major challenge has been to better understand its spread in order to effectively design interventions to combat it. In a recent work, Joseph B. Bak-Coleman and colleagues developed a generative model for misinformation engagement, which is defined as the total discussion and sharing of posts related to false information, and used this model to explore the efficacy of interventions to curb misinformation spread.

The model of misinformation spread has been equated to that of a viral dynamic spread within a population. At each time step, whether a post is made about a topic is determined by the virality in the previous time step. Once a user makes a post, the virality of that incident increases proportionately to that user's number of followers. As time proceeds, network saturation and competition for attention with other topics can reduce virality, which has been implemented in the statistical model by means of a decay function. To test the model, the authors collected data from Twitter during the 2020 US election period, as well as seeded the model with the estimated parameters, posts in the initial five-minute time step, and the empirical distribution of follower counts for each five-minute interval. The observed engagement for events showed a good agreement with the patterns observed using the model.

The authors tested various interventions such as removing or hiding all posts for a particular set of search terms, and what they called 'virality circuit breakers', which did not remove content but limited the spread of each post by means of different mechanisms such as suspending algorithmic amplification. All of these interventions, however, take time to be implemented, since the posts with false information need to be first identified. Alternatively, the authors also explored nudging, where the user is reminded to consider accuracy when sharing a post. Finally, the authors tested more extreme measures such as banning accounts that have been found sharing misinformation three times or more. While all of these interventions showed a decrease in the spread of misinformation, they can be very drastic, and for some of them, such as removing posts or banning accounts, there would be a time delay. To combat this, the authors recommended combining these interventions with milder implementations and, surprisingly, a modest implementation of these interventions led to a large reduction in the spread of misinformation.

Ananya Rastogi

Published online: 19 July 2022
https://doi.org/10.1038/s43588-022-00290-5