A Complex, Integrative Agent-Based Model of Disinformation Cascades

No Author Given

No Institute Given

Disinformation, or false information that is spread with the intent to deceive an audience, has been the subject of much social scientific study in recent years. From political information during a presidential campaign to the recent Covid-19 pandemic , false information has been found to spread farther and faster over a social network than true information.

We have developed a complex, integrative agent-based model of disinformation, the chief purpose of which is to evaluate the importance and sensitivity of different model components for replicating the spreading dynamics of disinformation on social media. Agent-based models of disinformation cascades are ubiquitous, with many practitioners using these models to predict the reach of a disinformation message, or identify a key network configuration or psychological feature that impacts spreading behavior. However, few modeling techniques integrate social and psychological factors together, along with a competitive information environment in which many messages compete for a person's limited attention, to demonstrate the nuances of complex interactions and the generalizability of modeling approaches to real-world disinformation cascades.

Along with varying network structures and information competition, our model integrates four social-psychological processes which have been found to influence spreading dynamics: trust, centrality, ideological consistency, and information accuracy. Trust is a directed edge characteristic (i.e., a receiver (dis)trusts a sender) that adjusts the message's 'believability', and thus whether an agent adopts the message. Network centrality impacts the message ordering in the receiver's inbox, much as a social media algorithm would promote messages from more central senders in a user's feed. Ideological consistency argues that people are more likely to adopt messages which comport with the pre-existing opinions which, toghether, shape their overall worldview. Our model utilizes multiple topics and agents' opinions on those subjects correlate to varying degrees. The message's difference from that opinion space determines adoption rates. Finally, information accuracy concerns the objective truth value of a message and whether, through subject knowledge or fact-checking behavior, a person ascertains that a piece of disinformation is false, and is thus less likely to forward the message to others. In our model, each message has an objective truth value and false messages are identified and rejected by agents with a probability that is dependent upon their subject knowledge. We present model results across a wide variety of feature settings and demonstrate a tool for identifying simulations which more closely replicate real-world cascade distributions.