2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

An Investigation into the Sensitivity of Social Opinion Networks to Heterogeneous Goals and Preferences

Patrick Shepherd, Mia Weaver, and Judy Goldsmith
Department of Computer Science, University of Kentucky, Lexington, KY, 40506-0633, USA
Email: patrick.shepherd@uky.edu and mia.weaver@uky.edu and goldsmit@cs.uky.edu

***Abstract*—As research into the dynamics and properties of opinion diffusion on social networks has increased, so too has the attention paid to modeling such systems. Simulations using *agent-based modeling* (ABM) analyze aggregate network outcomes when individual agents act on typically limited information, and tend to focus on agents that are conforming and homophilic — that is, they prefer to be around similar others, and they update their own personal state over time to be more like their friends. In this work, we illustrate the value of diverse agent modeling in environments that allow for strategic unfriending. We focus on network dynamics generated by three agent models, or *archetypes*. Our work shows that polarization and consensus dynamics, as well as topological clustering effects, may rely more than previously known on the interplay between individuals' goals for the composition of their neighborhood's opinions.**

I. Introduction

Current events have brought significant attention to the ability of online social interaction to effect outcomes in the real world. It is postulated that bad actors were able to sway the direction of both the Brexit vote in the UK and the US presidential election of 2016 (e.g., [18]). Even outside the realm of major historical events, discussion abounds of the effects — both good and bad — of selective exposure to information in online social networks, running the gamut from self-realization to radicalization. Whatever the true extent of these effects is, it has become hard to deny that they exist.

One of the major difficulties in developing robust models for such processes is the heterogeneity of individuals in the networks at hand. For instance, a simulation of political attitudes has to take into account the fact that agents will fall into one or more of a set of categories, and will act according to different motivations based in part on their group membership. This is not restricted to similarity-seeking agents, but much of the foundational and ongoing work in this vein focuses on agents that are more-or-less indistinguishable from one another in both goals and behaviors. Mostly, these studies are concerned with the emergence and stability of, e.g., opinion consensus within a network, and how to calculate optimal policies to bring about (or disrupt) consensus.

Research on diffusion and other processes over social networks considers some condition spreading through the network in discrete time steps, given a single rule to determine each node's state at each time step based in some way on the states of one or more of the node's neighbors. Research building on these models added features including: social influence, making some nodes more effective at propagating their state than others [2], [5], [9], [21]; masks, enabling nodes to either broadcast their state to the network or keep it hidden [11]; multidimensional diffusion spaces, which allow more than one phenomenon to spread through the network simultaneously [3], [5], [9], [20]; and in the multidimensional case, correlations between issues so that a node is more or less likely to flip its state in one dimension depending on its state in another [6], [9]. Most research in opinion diffusion rests on the assumption that all agents in a network prefer connections to others who hold the same opinion(s); there are exceptions to this rule, and some forms of opinion antagonism and their effects on the topology and opinion space of social networks have occasionally been investigated [10], [12], [13], [16], [19].

In this work, we examine the characteristic behavior of networks in the presence of three agent *archetypes*: homogeneous (HOM), heterogeneous (HET), and adversarial (ADV). These agent types are defined along two dimensions: *homophily (heterophily)* — the preference for connections to others who have similar (dissimilar) opinions — and *conformity (contrarian)* — the tendency to move towards (away from) friends' opinions. We explore a) how different mixtures of these agent types affect both the topology and opinion space of the network, and b) how resistance to opinion influence alters long-term network outcomes. This work illustrates the impact of multiple interacting agent types on a network, and opens the door for new research into the finer implications of these modeling strategies.

II. Related Work

Pilditch *et al.* [17] developed an agent-based opinion cascade model in which opinions diffuse through a network based on individual agent decisions. The authors develop synthetic networks in a unidimensional opinion space, with agents moving through opinion space based on their observations of their friends' known opinions. Agents observe the opinion of the first of their neighbors to declare it publicly, and use a simple Q-learning model to determine whether or not they will update their own opinions. Their experiments on a uniform agent type showed that the network stayed nearly evenly split in terms of opinions, but that clustering with like-minded agents was a dependable outcome.

IEEE/ACM ASONAM 2020, December 7-10, 2020
978-1-7281-1056-1/20/$31.00 © 2020 IEEE

Duggins [8] provides a similar model. In it, agents have attributes for tolerance of dissimilar opinion, susceptibility to social influence, and desire to conform to social norms, where "conformity" refers to how far between an agent's true opinion and the socially normative opinion the agent wishes to appear to others. Simulations involve individuals in neighborhoods expressing an opinion somewhere between what they truly believe and the socially normative opinion. Ye *et al.* [22] also investigate a model featuring separate private and public opinions for agents. Madsen *et al.* [15] use an update scheme roughly the same as the bounded confidence model [4]. Agents have a real-valued opinion about the state of the world. Each time step agents seek out others whose opinion is close enough to their own, then update their opinion based on the aggregation of observed opinions.

Chen *et al.* [3] model agent-based opinion dynamics with an additional personality parameter for modeling homophily. New edges are created between pairs of nodes with probability proportional to the Euclidean distance between their opinion vectors. Li *et al.* [14] use the Stubborn Individuals and Orators (SO) model [7], which uses two additional parameters to model how resistant individuals are to opinion change, and how influential individuals are, both of which our model accommodates. Banisch *et al.* [1] explore the dynamics of opinion formation when social feedback is used as a reinforcement learning signal.

III. The Model

We begin with a traditional network graph $G = \langle V, E \rangle$, where V is the set of *nodes*, or *agents*, and E is the set of *edges*, or connections between them. Many settings require that a weight vector W be included in the definition of the graph, but for this work we assume that all edge weights are equal. Each node is embedded within a k-dimensional binary opinion space, and their positions in this space change over time dependent on the positions of adjacent nodes. We refer to node i's position in opinion space at time t as $\vec{b}_i^t = \{-1, 1\}^k$. We will refer to i's opinion on topic k at time t as b_{ik}^t. Each agent in our networks embodies one of three archetypes that govern their behavior.

Archetypes are defined by properties relating to preferences over opinions as well as opinion mobility. The HOM archetype is both *homophilic* and *conforming* — i.e., HOM agents prefer connections to other agents with similar opinions (the more similar the better), and over time they move closer to the majority opinion in their neighborhood. ADV agents are the opposite in both ways; they are *heterophilic* and *contrarian* — i.e., they prefer connections to others who disagree with them, and they move away from their neighborhood's majority opinion. These two types are reasonable proxies for people who prefer similarity — a widely known human characteristic — and people who prefer to delineate themselves from the crowd. Our final archetype, HET, is meant to represent another human cohort: people who do tend to conform over time to the majority opinion, but prefer a wide array of opinions within their neighborhood. Functionally, the implementation of each archetype is achieved through the combination of an *update rule* and a *reward function*.

The update rule for each agent is a function taking a neighborhood's average opinion as input and producing a new opinion vector. An update rule is of the form: **if** *neighborhood average opinion is strong enough* **and** *opposite of what I want,*[1] **then** *flip my opinion with some probability*. The threshold for "strong enough" and the definition of "what I want" are design decisions. At each time step, agents see their neighbors' opinions and use those to update their beliefs. When agent i goes to update its opinions, it first calculates the average opinion on each topic k within its neighborhood as $\widetilde{b}_{ik} = |N(i)|^{-1} \sum_{j \in N(i)} b_{jk}$. The agent can then determine whether its neighborhood mostly agrees with it by testing $\widetilde{b}_{ik} * b_{ik} < 0$ or not. This information can then be used to determine how i's opinions move through opinion space.

Conforming archetypes update their opinions to move toward their neighborhood majority opinion. Thus, if agent i is either HOM or HET, its opinion update takes the following form:

$$b_{ik}^{t+1} = \begin{cases} b_{ik}^t & \text{if } \widetilde{b}_{ik} * b_{ik} \geq 0 \\ -b_{ik}^t & \text{if } \widetilde{b}_{ik} * b_{ik} < 0. \end{cases}$$

Reversing the signs on the resulting cases gives the update rule for contrarian agents, who move away from the majority opinion. In the case that an agent decides to flip its opinion on a topic, the flip happens with probability 0.5 for our experiments.

The reward function for each archetype must describe its preferences for neighbors' opinion profiles relative to its own. HOM agents prefer connections to others with similar opinions, so their reward function should correlate positively with the amount of agreement they have with their neighbors to reflect this. The simplest way to accomplish this is to set their reward function $r_i^{hom}(j) = 1 - d(i, j)$, where $d(i, j)$ is the percentage of disagreeing entries between b_i^t and b_j^t. Similarly, we can define the reward function for ADV agents as $r_i^{adv}(j) = d(i, j)$, which is maximized with neighbors who agree with the agent on nothing at all. Finally, HET agents' reward function is set to be maximized by a neighbor who agrees on exactly half of the topics, and disagrees on the other half, decreasing linearly to 0 as the neighbor tends toward total agreement or disagreement.

Agents are also given *policies* based on their archetypes. These policies are simple constructions to allow agents to sever connections they consider not worthwhile in terms of reward. This allows all agents in our networks to self-organize to fit their preferences. A final aspect we associate with archetypes for this work is resistance to influence. This is a scalar parameter determining each agent's threshold for opinion change — more resistant agents will not consider changing an opinion unless a more significant portion of their neighbors have the desired opinion.

[1] We use the term "want" to reflect that some agents do not desire agreement with the majority, but rather prefer that the majority holds the opinion opposite their own.

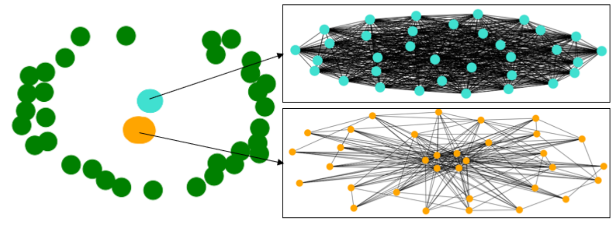

Fig. 1: Fully mixed networks with uniform resistance to opinion influence separate into disjoint groups. Types are HOM (blue), HET (green), and ADV (orange). Node positions are arbitrary, but more densely connected clusters are shown physically closer than sparsely connected clusters.

IV. Experiments

We conducted several experiments to investigate the effects of different agent types on the evolution of a network's topology and opinion space. We illustrate these archetypes' effect on the behavior of networks they inhabit. In particular, we investigate: a) the effects of different proportional mixes of agent types, b) initial network densities, and c) resistance levels by agent type. All networks used were of 100 agents, and were small world networks with average degree dependent on the experiment.

1) The Effects of Network Composition: To investigate the impact of the agent type distribution on network outcomes, we ran 10 100-step simulations on networks with each of the following distributions (% HOM / % HET / %ADV): 34/33/33, 50/25/25, 60/20/20, and 70/15/15. For these experiments, we held agent resistance to 0 for all agents.

Figure 1 shows a typical outcome in these networks, independent of type distribution. The network in the figure resulted from a 70/15/15 run, but the apparent patterns existed in all conditions tested. Regardless of the distribution of agent types, these networks almost always separated into three cohorts: the HET agents, who end up isolated in the network as consensus begins to take over in the core; the HOM agents, who again aggregate into a complete subgraph (unless their numbers were great enough, in which case they typically split into two disjoint, disagreeing clusters); and the ADV agents form a core periphery cluster. Further, in each simulation the network would remain without isolates for several steps until one HET agent left; once that happened, the rest of the HET agents left very quickly thereafter. The ADV and HOM clusters, though already formed, never separate from each other until most or all of the HET agents leave.

2) The Effects of Initial Density: The initial density of the network can have significant influence over its evolution. For example, in pure ADV networks an initial density set too small will cause the network to fragment more. We used the same conditions as in the previous experiments, but with average degree $= 5, 10, 15, 20,$ and 25. At average degree 5, agents did not have enough connections to form anything more than two- to six-member components under any type distribution. However, more evenly-split networks did tend to break apart

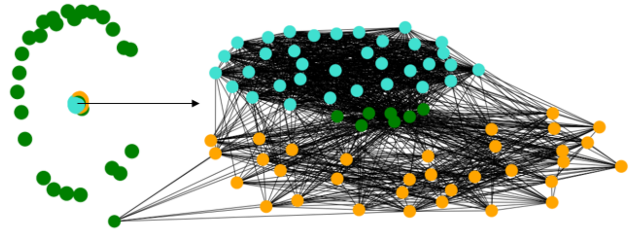

Fig. 2: When HET agents have a higher resistance to influence, they prevent networks from separating into disjoint camps.

into smaller groups on average. ADV agents also appeared to benefit from this slightly, in some cases maintaining a cluster with one of the other agent types.

An overly sparse initial network lends itself to fragmentation as would be expected, whereas sufficiently dense starting networks appear to almost always separate into disjoint components delineated by agent type. The novel illustration is that *polarization is not restricted to opinion space*. In other words, in our experiments we saw not only the expected polarization in opinion, but also in the proclivity of agents to self-segregate based on type. Even HET agents, before they split apart entirely, formed tighter communities with each other than with either of the other archetypes. This means that even non-homophilic agents tend to bond most strongly and persistently with each other.

3) The Effects of HET Agent Resistance Levels: The observations of our last two sets of experiments make clear that agents seeking balance may have additional complications finding a suitable situation for themselves within the network given the behavior of other archetypes. These agents also seem to have a cohesive effect on the network as a whole. Whenever mixed networks split apart along opinion and/or archetype lines, they only do so after most of the HET agents have left.

It makes sense that these agents would choose to leave the network once the other types have entrenched themselves in their own segregated camps, because by definition HET agents have two reward "valleys": total agreement and total disagreement. These two valleys overlap the reward peaks of the other two archetypes, creating a balancing act between them. Our last set of experiments is designed to test the aggregate effects of endowing HET agents with greater resistance to opinion influence.

Agents in this style environment who have a greater openness to different opinions tend to foster consensus rather than hinder it [16]. We varied HET agents' resistance value from 0.0 to 0.5, which corresponds to agents needing between half and 3/4 of their neighbors to disagree with them on a topic before they might flip their opinion.

In most tests with resistance set to 0.0 for HET agents, the networks split apart into camps. Figure 2 illustrates the characteristic outcome we observed when we increased that value to 0.25. After 100 steps, most of the HET agents in the network had left, just as before. However, a small cluster of them remained in between the two other clusters, which

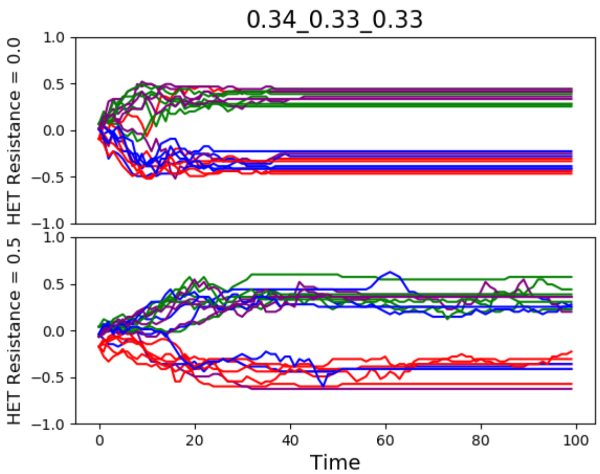

Fig. 3: The effect of resistant HET agents on opinion consensus. Both plots show the network-wide average opinion on each topic, where the x-axis shows time, and the y-axis shows the average opinion. Each color corresponds to one topic, and multiple lines of the same color represent different runs begun from the same initial conditions. If a line rises all the way to $y = 1$, it means every node in the network adopted an opinion of 1 on that topic. If a line stays at $y = 0$, then half the nodes have opinion 1 and the other half have opinion -1.

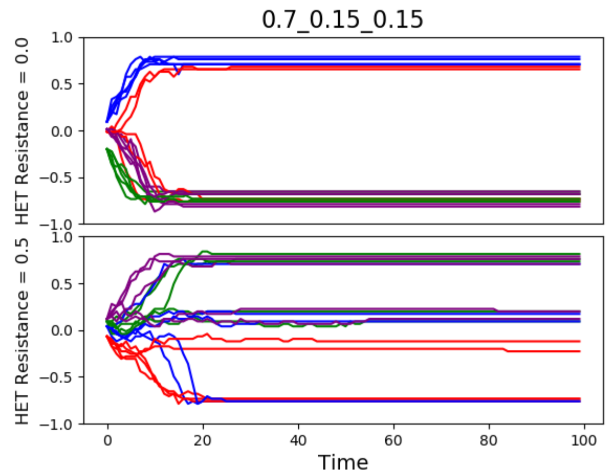

Fig. 4: HET agents can have an even larger effect when HOM agents are a significant majority. Resistant agents help network-wide opinions to maintain multiple well-represented sides. When these agents are not resistant to opinion influence, near-complete consensus manifests across all topics.

organized themselves in their characteristic ways. It can be seen that one agent was possibly about to leave the network, even after 100 steps, so it is immediately evident that HET agents alleviate some topological rigidity.

A. *Effects on Consensus*

It has been shown that the presence of heterophilous agents in a network help foster consensus [16]; our findings support this idea. Figure 3 shows the movement of opinion averages over 100 steps in a network evenly split between all three archetypes. The panels correspond to two levels of HET agent resistance to influence: 0.0 (top) and 0.5 (bottom). Each opinion has its own color, and lines of a single color represent the average opinion on a single topic across the network. Each line shows the average of a single opinion over a separate experiment from the same initial conditions. The figure shows that, when HET agents probabilistically update their opinion whenever a strict majority of their friends have the opposite opinion (i.e. HET agents had a resistance level of 0.0, top panel), the network settles into a stable opinion configuration at around $t = 40$. Here, the network exhibits a roughly 2/3 to 1/3 weighting on opinions of 1 versus opinions of -1 for every topic. Alternatively, when HET agents require more disagreement from their friends with their own opinions to consider changing (bottom panel), stable opinion configurations often do not emerge at all. The network still seems to find a relative "comfort zone" with respect the distribution of opinions, but each step sees some individuals change their mind. This seems to show that having a cohort of relatively stubborn HET agents does not generally cause chaos in the opinion space of a network, but rather fosters its ability to remain somewhat fluid.

Figure 4 shows a more dramatic effect. Again, the top panel represents a network with HET agents who update their opinion probabilistically based on the majority rule. It can be seen that, on every topic, almost all individuals in the network had the same opinion — either 1 or -1 — i.e., there were no topics that had a significant number of proponents on both sides. However, when HET agents were more resistant to changing their mind, several topics remained approximately evenly split; about half the network had an opinion of 1, and the other half had an opinion of -1. Even though the resistant HET agents made up only 15% of the network, their presence caused a severe disruption in the process of average opinions drifting to the two opinion poles. These results together support the idea that heterophilic agents who follow the same update rules as their homophilic counterparts tend to lead networks toward opinion stability.

Once those configurations were established, there was no further deviation. Notable, then, is the fact that resistant HET agents seem to foster opinion fluidity in the network, but not radically. That is, their presence, even in small numbers, appears to be associated with an opinion space that quickly settles into a general configuration within which some individuals are constantly exploring; it is not the case that these networks exhibit highly chaotic evolutions in opinion over time. We ran the same set of experiments again, but with a time horizon of 300 steps to see if these networks did eventually settle into a stable configuration. We found that they did not. Networks with some portion of influence-resistant HET agents are able to maintain a relatively, but not completely, stable opinion configuration across longer spans of time than networks without. With resistant HET agents present, HOM agents no longer formed a unanimous opinion set across all topics. There was still consensus on some topics, but others found a stable configuration with roughly half the agents representing each of the two opinions (1 or -1). The HET agents were also unanimous on some topics, although they showed more variability. Finally, the ADV agents did not all reach a unanimous position on any topic, and showed the most variability in terms of opinion space outcomes.

These results illustrate the extent to which network dynamics can be influenced by the presence of different agent types, and begs further investigation into more archetypes. Specifically, while traditional conforming homophilic and contrarian heterophilic agents influence the network in predictable ways, the addition of conforming agents that are directly in the middle of the homophilic/heterophilic continuum and are harder to influence produced unexpected outcomes. They introduced a new topological dynamic due to their particular behavior, and also fostered fluidity and dissuaded unanimity in opinion space.

V. Conclusion

In this work, we modeled different agent archetypes for opinion diffusion simulations, defined both by their notion of utility as well as their behavior during the state-update phase of each time step. We then presented results from an investigation into the interactions of some different types.

We explored the collective behavior of typical homophilic agents who become more like their neighbors, heterophilic agents who are contrarians, and agents who try to balance similarity with dissimilarity in their own regions of the network. Our tests showed that, when HET agents are present in the network, they often act as bonds that hold the network mostly together, keeping self-segregated clusters of HOM and ADV agents from becoming completely disconnected from each other. We observed repeatedly that networks tended to stay connected until the ejection of a HET agent, which seemed to have a cascading effect ending when the last HET agent separated and the clusters they were connecting come apart entirely. Networks with mixed agent types appear to be sensitive to multiple conditions, of which we studied the amount of resistance HET agents have to external influence, and the relative proportions of the archetypes in the network. The amount of HET resistance played a pivotal part in determining network outcomes with respect to those agents. In opinion space, the parameters of HET agents were also important. Networks with a more even mix of agent types permitted a more even representation of opinions on all topics — that is, the network was split in approximately a 2:1 ratio of opinions of 1 versus -1 or vice versa. Networks that had fewer HET and ADV agents were prone to converge to a nearly network-wide consensus on all topics. Resistant HET agents appeared to alleviate this condition.

This work lends itself to several extensions. More robust methods for creating new links must be implemented to more closely mirror the mechanics of real-world networks. We believe that this dynamic will uncover further information about long term network configurations. Further, the notion of reward for HET agents is currently rudimentary, and more nuanced alternatives need to be explored. Finally, the possibilities for learning agents could have broad real-world implications.

References

[1] Sven Banisch and Eckehard Olbrich. Opinion polarization by learning from social feedback. *The Journal of Mathematical Sociology*, 43(2):76–103, 2019.

[2] Fabian Baumann, Philipp Lorenz-Spreen, Igor M Sokolov, and Michele Starnini. Modeling echo chambers and polarization dynamics in social networks. *Physical Review Letters*, 124(4):048301, 2020.

[3] Xi Chen, Jie Tang, and Yizhou Sun. Modeling personalized dynamics of social network and opinion at individual level. *arXiv preprint arXiv:1811.02791*, 2018.

[4] Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(01n04):87–98, 2000.

[5] Morris H. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[6] Daniel DellaPosta, Yongren Shi, and Michael Macy. Why do liberals drink lattes? *American Journal of Sociology*, 120(5):1473–1511, 2015.

[7] Alessandro Di Mare and Vito Latora. Opinion formation models based on game theory. *International Journal of Modern Physics C*, 18(09):1377–1395, 2007.

[8] Peter Duggins. A psychologically-motivated model of opinion change with applications to American politics. *arXiv preprint arXiv:1406.7770*, 2014.

[9] Noah E Friedkin, Anton V Proskurnikov, Roberto Tempo, and Sergey E Parsegov. Network science on belief system dynamics under logic constraints. *Science*, 354(6310):321–326, 2016.

[10] Serge Galam and Taksu Cheon. Asymmetric contrarians in opinion dynamics. *Entropy*, 22(1):25, 2020.

[11] Umberto Grandi, Emiliano Lorini, and Laurent Perrussel. Strategic disclosure of opinions on a social network. *arXiv preprint arXiv:1602.02710*, 2016.

[12] Frans Jacobs and Serge Galam. Two-opinions-dynamics generated by inflexibles and non-contrarian and contrarian floaters. *Advances in Complex Systems*, 22(04):1950008, 2019.

[13] Evguenii Kurmyshev, Héctor A Juárez, and Ricardo A González-Silva. Dynamics of bounded confidence opinion in heterogeneous social networks: Concord against partial antagonism. *Physica A: Statistical Mechanics and its Applications*, 390(16):2945–2955, 2011.

[14] Ying Li, Hongduo Cao, and Genfeng Wen. Simulation study on opinion formation models of heterogeneous agents based on game theory and complex networks. *Simulation*, 93(11):899–919, 2017.

[15] Jens Koed Madsen, Richard M Bailey, and Toby D Pilditch. Large networks of rational agents form persistent echo chambers. *Scientific Reports*, 8(1):1–8, 2018.

[16] Sebastien Motsch and Eitan Tadmor. Heterophilious dynamics enhances consensus. *SIAM Review*, 56(4):577–621, 2014.

[17] Toby D Pilditch. Opinion cascades and echo-chambers in online networks: A proof of concept agent-based model. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society (CogSci 2017)*. Cognitive Science Society, 2017.

[18] David E Sanger and Catie Edmondson. Russia targeted election systems in all 50 states, report finds. *New York Times*, July 25, 2019.

[19] Alina Sîrbu, Vittorio Loreto, Vito DP Servedio, and Francesca Tria. Opinion dynamics with disagreement and modulated information. *Journal of Statistical Physics*, 151(1-2):218–237, 2013.

[20] Giuseppe Toscani, Andrea Tosin, and Mattia Zanella. Opinion modeling on social media and marketing aspects. *Physical Review E*, 98(2):022315, 2018.

[21] Gérard Weisbuch, Guillaume Deffuant, Frédéric Amblard, and Jean-Pierre Nadal. Meet, discuss, and segregate! *Complexity*, 7(3):55–63, 2002.

[22] Mengbin Ye, Yuzhen Qin, Alain Govaert, Brian DO Anderson, and Ming Cao. An influence network model to study discrepancies in expressed and private opinions. *Automatica*, 107:371–381, 2019.