Automatic Control of Opinion Dynamics in Social Networks

Michael DeBuse and Sean Warnick†

Abstract— **This work addresses the issue of how the common interaction dynamics of social media networks enable the creation of "echo chambers," or self-reinforcing, disjoint communities with distinct opinion biases. We theorize a Strategic Agent as a feedback controller that can dismantle echo chambers and encourage an overall healthier sharing of opinions among agents in the network. We then show how this same controller can then be used to drive opinions of all agents on the network to any desired opinion bias, emphasizing the importance of ethical use of automatable discourse-building and enabling technologies, such as large language models like Chat-GPT.**

I. INTRODUCTION

Social media has created a new landscape for sharing and shaping opinions through posted content. In this environment, we observe that people tend to selectively follow like-minded individuals while often disregarding those with opposing views [1]. This follow/unfollow mechanism results in opinions diffusing through a dynamically changing network. Social media users with strong and similar opinions often gather to reinforce their beliefs, while excluding dissenting voices. This isolationist behavior, known as an Echo Chamber [2], fosters misinformation and societal polarization. The World Economic Forum's Global Risks Report 2023 highlights these as significant concerns [3].

This research recreates the tendency towards Echo Chambers in a social network and then investigates the feasibility of a Strategic Agent in limiting this societal polarization by successfully rebuilding healthy discourse between formerly opposing opinion groups. First, we attempt to rebuild conversation between isolated, biased opinion groups by moderating the extremity of those biases, making them more tolerable. Next, we show that this same Strategic Agent can under reasonable assumptions drive a social network toward any desired opinion bias, thus becoming a controller of opinion over the social media network. We end by emphasising the importance of ethical use of automatable discourse-building and enabling technologies, such as large language models like Chat-GPT, that can be used to drive opinions in a similar manner.

II. RELATED WORKS

Social networks [4] as a platform for diffusion model analysis have been highly studied from adopt-or-not models [5], multi-agent models [6], to models that investigate the effects of uncertainty [7] or agent profile data [8], [9] on the rate and behavior of the diffusion. They have garnered a lot of attention from the control communities for the unique dynamics they enable [10], [11], especially in modeling the affect of various means of influencing spread dynamics.

The Taylor Model [12] implements immutable, external communication sources in the network to influence those vertices connected to them. The Friedkin-Johnsen model (FJ model) [13] is a discrete-time variant that introduces agent stubbornness as a means of opinion control. Roberk Bredereck and Edith Elkind address the time-varying nature of social network structure through selective edge modification and show that diffusion is still possible, but their methods are computationally hard [14]. Research by Haibo Hu [15] studies the effects both of stubborn (committed) opinion holders and external influences on opinion. Similarly, Heather Brooks and Mason Porter show that external media can directly influence the ideology of content shared in social networks through those agents following the social media accounts [16]. Wendt et. al. show how a perfectly stubborn agent can drive the opinions of a network to a desired value despite agent resistance to influence. In each of these cases, the controller agent is either external or unchanging in it's opinion. Our Strategic Agent is a feedback controller with the ability to move throughout the network and modify its opinion as needed to reach its goal.

Diffusing opinions over social networks can be seen as a form of consensus of that opinion over the agents of the network. Consensus over social networks is another widely researched field, from group decision making [17] to the affects of variable trust in that decision making over the network [18]. Acemoglu et. al. show how the presence of immutable agents in a social network prevent global consensus of a single opinion [19]. Echo chambers can similarly act as immutable opinion sources the prevent consensus. The creation and effects of echo chambers have been studied using opinion thresholding [20], [21], limited communication resources [22], and repulsion edges [23]. Baumann et. al. show that such echo chambers reliably appear using politically polarizing debates from Twitter [24]. Our research utilizes thresholding in the tolerance of opinion difference to recreate the tendancy towards echo chambers that prevent consensus. We then show how consensus can still be obtained despite the presence of echo chambers through the use of our Strategic Agent.

III. PROBLEM FORMULATION

Our model for strategically influencing the biases in a social network consists of three pieces: the social network model, the social network dynamics, and the dynamics of a proposed *Strategic Agent* who acts as a feedback controller

†Michael DeBuse and Sean Warnick are with the Information and Decision Algorithms Laboratories (IDeA Labs), Department of Computer Science, Brigham Young University, Provo, UT 84602, USA mdebuse3@gmail.com, and sean@cs.byu.edu

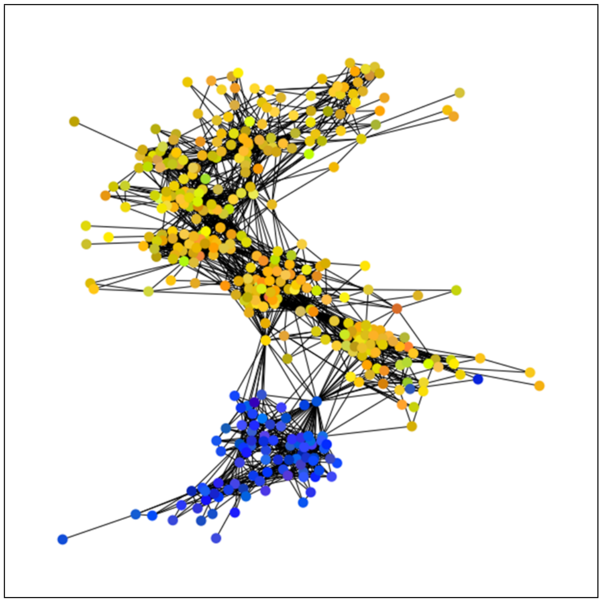

Fig. 1. Initial state of the undirected Infect-Dublin graph with the two major ideological groups represented by yellow and blue.

in the system, responsible to drive the biases throughout the network to desired values.

A. SOCIAL NETWORK MODEL: AGENTS AND BIAS

Our model of a social network is an undirected graph, $G = \{V, E\}$, where the vertex set V is a list of N *agents*, and edges are ordered pairs of vertices or nodes. An edge from node $v_j \in V$ to $v_k \in V$ indicates that agent k is *following* agent j and thus receives communication from agent j. Looking the other direction, agent j is an *influencer* of agent k if agent k is a *follower* of agent j. The set of influencers or neighbors of an agent k, $\mathcal{N}^k$, is the set of agents j such that an edge $(v_j, v_k) \in E$.

We associate with each agent j a bias, $x^j \in \mathbb{R}^n$, which is a vector quantified to represent the opinions of agent j on n distinct topics. The element, x^j_i, then models agent j's opinion on topic i. In this work we restrict the range of elements in any opinion or bias vector x to be within the interval $[0\ \ 1]$, with 0.5 indicating a completely neutral stance on the topic.

B. SOCIAL NETWORK DYNAMICS

In this work we consider two actions available to any agent at any given time. The first is to request or to accept a request to follow another agent. If agent k requests to follow agent j, and if agent j accepts (acceptance probability is 1.0 for this study), then agents j and k will follow each other. Mutual following allows us to use undirected graphs to represent social networks, as shown in Figure 1. These requests and disconnects (explained later) change the underlying structure of the network, the edge set E in the graph G. In this paper we randomly choose a number of agents, s, in the network to each solicit one other agent, and s is chosen to roughly balance the number of cancellations so the overall number of connections in the graph remains relatively constant.

The second action an agent may take is to publish a social media post on their account, viewable only to those agents following the posting agent. To model the post's impact on agent opinions, we say that agent j publishes a post at time t, and the topical content of the post is represented by a vector $u^j(t) \in \mathbb{R}^n$ for our n modeled topics. Elements $u^j_i(t)$ of this vector may take values of 0.5 on topics i that are irrelevant to the post (i.e. the post is neutral on those topics).

To model how posted content affects followers' opinions, we assume full-state feedback (inspired after public-facing profile information), meaning that $x^j(t)$ is known for any agent in the network, $j = 1, ..., N$. People's opinions move *as a collection* towards those of other people they agree with on some particular topic. That is to say, we are exploring the situation where agreement on at least one topic establishes *trust* between agents, and this trust causes the following agent to adjust their *entire opinion vector* in a direction towards that of the influencing agent. Note that we are not suggesting that the following agents adjust their opinions in the direction of the *posted content*, as one might assume would happen if the posted content were an effective logical argument capable of changing one's mind on the issue. On the contrary, here we are modeling an entirely different situation, where *opinions change in directions mimicking the public facing opinions of producers of pleasing or agreeable content*, regardless of the sensibility or logical correctness or veracity of a particular post. We make this distinction to more directly model shifts in societal polarization based on internal opinion bias, as stated in the introduction.

To characterize these effects, we define two thresholds, the *trust* threshold, $0 \le \theta_r \le 1$, and the *distrust* threshold, $0 \le \theta_d \le 1$, where $\theta_r < \theta_d$. Posts within the trust threshold of an opinion vector cause the opinion vector to move in the direction of the author of the post (see Equation (1)), while posts outside the distrust threshold of an opinion vector cause the following agent to disconnect from the posting agent. We also consider an $N \times N$ (row stochastic) susceptibility matrix, A, where $0 \le a_{jk} \le 1$, $\sum_k a_{jk} = 1$ for all j, and $a_{jj} \neq 0$ for all j, reflects the relative susceptibility of agent j to be influenced by agent k. These allows us to model non homogeneous agents, with some being much more stubborn and resistant to influence—at least from certain agents—than others.

The basic dynamics for the social network then proceed as follows. At each time instant t, all agents post new content. If an agent k receives content similar enough to its own opinion on a related topic to be within the trust threshold, then the indices of all influencer agents $j \in \mathcal{N}^k(t)$ posting such pleasing content at time t are included in a set $\mathcal{J}^k(t)$. That is to say,

$$\begin{aligned} \mathcal{J}^k(t) \quad = \quad & \{j \in \mathcal{N}^k(t) |\ \exists\, i \text{ such that} \\ & |u^j_i(t) - x^k_i(t)| \le \theta_r \text{ and } (v_j, v_k) \in E\} \end{aligned}$$

Related to this set of influencers on agent k at time t is the *active susceptibility*, $\alpha^k(t)$, which is the sum of the susceptibilities, a_{kj}, for all $j \in \mathcal{J}^k(t)$. With this set of influencers on agent k characterized, then agent k updates

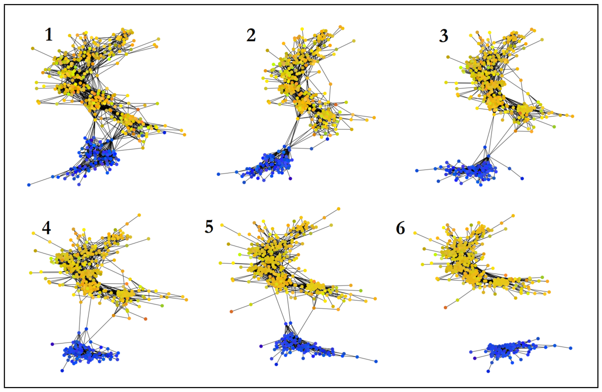

Fig. 2. Progression of the Infect-Dublin network towards disparate partitions over time. Node color indicates agent opinion orientation or bias. Random requests to follow, creating new edges, are balanced by unfollowing, when influencer posts are sufficiently different from a follower's current bias. Notice how initial variation in bias is eliminated in echo chambers (one yellow, one blue), and these distinct communities eventually completely disconnect from each other while self-reinforcing a dominant orientation (or particular shade of yellow and blue) within the group.

its opinion or bias vector as follows:

$$x^k(t+1) = \sum_{j \in \mathcal{J}^k(t)} a_{kj} x^j(t) + (1 - \alpha^k(t)) x^k(t) \qquad (1)$$

These dynamics indicate that agent k updates its opinion vector to be the convex combination of its current opinion with the opinion vectors of agents that k follows who post content at time step t that is similar to k's current opinions on at least one topic.

When an agent k is offended by a recieved post, meaning $|u_i^j(t) - x_i^k(t)| \geq \theta_d$ for a topic i and an agent $j \in \mathcal{N}^k(t)$, they disconnect from agent j. Posts that fall between θ_r and θ_d have no effect on agent opinions or graph structure. Figure 2 shows how the combined model dynamics often result in strong "echo chambers" that eventually can completely disconnect into distinct subgraphs or communities.

C. *STRATEGIC AGENT AND FEEDBACK CONTROL*

The Strategic Agent is an automated agent within the social network (not an external influence) who acts as a controller that drives the opinion bias in the network towards some target bias. It has three tasks:

1) Strategically solicit agents to build its follower base,
2) Design posts with strategic sequences of content to move opinions of other agents in the network without triggering distrust,
3) Appease its followers so that it is not dropped by the unfollow rules by agents that are useful to its mission.

The Strategic Agent aims to drive opinions in the network by performing the three tasks mentioned, acting as an automated "socializer" to counteract echo chambers and encouraging agent exposure to diverse opinions instead of reinforcing existing beliefs.

To accomplish this goal, the Strategic Agent needs to both 1) plan and execute solicitations as well as carefully designed post sequences that give it *positionality* in the network that make the network *controllable* from its position, and 2) use its positionality to drive opinions in the desired manner. The Strategic Agent solicits s new agents at random per round where the edge distance $dist(v_{StrategicAgent}, v_{new}) \geq 2$ for the relevant vertices $v \in V$ in the network. Note that distance here is measured as the number of edges traversed between nodes. Identifying these s agents is easy on most social media platforms, since these platforms typically allow users to identify mutual friends. Any randomly selected agent with no mutual friends with the Strategic Agent is a viable candidate. In this way, the Strategic Agent can broaden its influence without the need to know in-depth knowledge of the structure of the social network. It only needs the means to search for accounts, another feature included in most social media platforms.

To prevent loss of followers and increase influence through the dynamics described in Equation 1, the Strategic Agent must set its opinion biases to the most tolerable values according to its neighboring agents. We can imagine that once the optimal opinion bias is known at the current time-step, a human or text-generative system like ChatGPT [25], that is near-indistinguishable from a human, could write a social media post matching the needed bias.

Since each follower of the Strategic Agent has its own individual trust interval of $T = [x_i - \theta_r \quad x_i + \theta_r]$ for each opinion (the distance of θ_r from the agent's x_i), setting the bias becomes a Maximum Overlapping Intervals problem. To set the bias for a single opinion, x_i, we implement the following algorithm:

1) Calculate T for the x_i of all neighbors of the Strategic Agent to give bias values with associated **start** and **stop** markers for every interval.
2) Place all individual bias values in a sorted list with their label of **start** and **stop**.
3) Iterate through list:
   a) If the marker is **start**, increment count. If it is the highest count so far, remember start value.
   b) If the marker is **stop**, decrement count. If count before decrement was the highest so far, remember stop value.
4) Return the start or stop value closest to the Strategic Agent's target bias. If the target bias is between the two, return the target bias.

We repeat this algorithm for each of the N different opinions.

IV. SIMULATION SETUP AND EXPERIMENTS

We us the Infect-Dublin network [26] as an example of a social network. We initialize the nodes, or agents, of the network with opinion vectors $x \in \mathbb{R}^3$ so that opinions can be visualised by RGB assignments corresponding to those three values (we will refer to the individual elements of x as x_R, x_G, and x_B from here on). As we explain at the end of Section III-A, strong differing opinions between agents in the network can create disjoint partitions of the network, or Echo Chambers. To enable this, we must initialize each individual opinion vector x^k with a different opinion bias. Using the

Girvan-Neuman method [27], we obtain two strong cliques within the Infect-Dublin network (the largest of which could be broken down further, but for the purposes of this research we keep the number of cliques as two). We then create templates for those two cliques, $[0.9, 0.75, 0.1]$ (yellow) and $[0.1, 0.25, 0.9]$ (blue). Agents are randomly initialized around the template using a Gaussian distribution with a variance of 0.1. This introduces diverse initial biases towards the three opinions to mimic real people. Values above 1 or below 0 are capped accordingly. The randomization ensures a range of opinions, including extreme and moderate to foster possible initial acceptance of differing opinions (as explained in Section III-B). The resulting initial state of the network can be seen back in Figure 1.

To facilitate a dynamically changing network given our two starting templates for each clique, we set $\theta_r = 0.25$ and $\theta_d = 0.75$. By having a low trust threshold, we create a challenge for the Strategic Agent to influence its followers. A high distrust threshold means that only those with strong opinions will cancel any follows from agents who create posts they dislike. By setting the initial bias templates to near-polar values of 0.1 and 0.9, we ensures that many members of each clique have a higher chance of disliking any post made by the opposing clique since $0.9 - 0.1 > \theta_d$. The only opinion they may tolerate initially is x_G since the difference of the two templates is between θ_r and θ_d. Again, this creates a challenge for the Strategic Agent, since in order to appease one clique, it must take a stance that is likely hated by the other, removing their trust. We initialize the susceptibility matrix, A, with the same susceptibility for each agent pair a_{kj}. Additionally, we weight the susceptibilities of influencing agents at each time-step such that any single agent will only be influenced by 25% the difference between its current bias and its influencing neighbors, meaning $\alpha^k \leq 0.25$. This treats α as the diffusion rate of the network.

So that the Strategic agent always has neighbors to try and influence, we set $s = 5$ for the Strategic Agent, meaning it solicits 5 new follower requests per time-step. We also set acceptance of the solicitation to 100%, but should the Strategic Agent make a post disliked by its new follower in the next time-step, it will immediately lose that follower. Essentially, by setting solicitation rate to 100% success, we are letting the Strategic Agent's decided post determine if any new followers will remain followers.

We perform three sets of experiments. In the first experiment setup, we attempt to prevent or fix echo chamber creation by driving the opinions of the agents on x_R, x_G, and x_B towards the neutral value of 0.5. To do this, we set the target bias for the Strategic Agent to $[0.5, 0.5, 0.5]$. Our assumption here is that by driving opinions towards 0.5, the difference between agent opinions on the three stances will be less than θ_d and have a higher chance of new follows appearing between the formerly disparate partitions.

The second experiment setup is to see if the Strategic Agent can directly drive all agents in the network towards any desired bias. We run three variations of this experiment using the target biases of $[0.9, 0.75, 0.1]$ matching the the

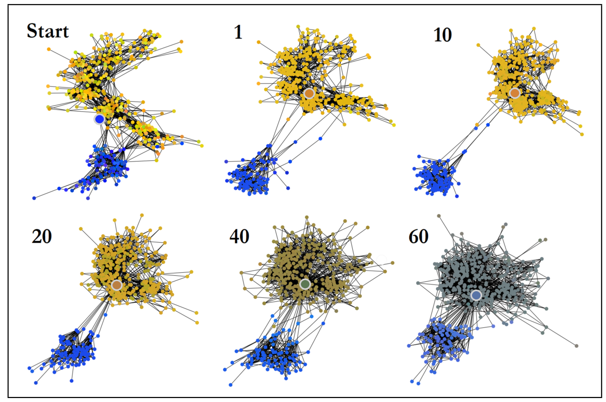

Fig. 3. Progression of the Infect-Dublin network from the first experiment with the Strategic Agent targeting a neutral bias of $[0.5, 0.5, 0.5]$. The Strategic Agent is represented by the larger vertex with the light grey border.

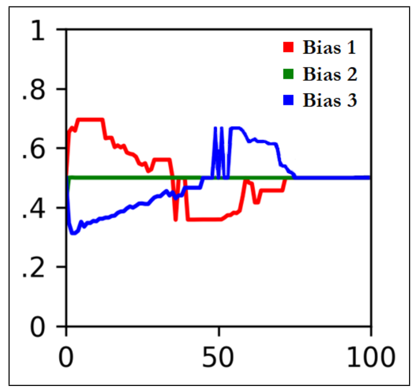

Fig. 4. Strategic Agent internal bias over time. It focuses on high red bias immediately to appease the yellow clique. At about time-step 40, it swaps to a low red and high blue bias to appeal to the blue clique. At time-step 75, all agents can tolerate neutral bias, and so the Strategic Agent sets its bias to neutral to drive all agents to neutral.

yellow bias template, $[0.1, 0.25, 0.9]$ matching the blue bias template, and $[0.4, 0.8, 0.4]$ for a new green bias template.

The final experiment setup is to have two different target biases for the Strategic Agent. The Strategic Agent starts driving the agents of the network towards neutral bias of $[0.5, 0.5, 0.5]$ and then changes to the extreme bias of the yellow clique, $[0.9, 0.75, 0.1]$. The reason for this experiment is due to the limitations of the second experiment that prevent the Strategic Agent from completely driving the network directly to some target biases.

V. RESULTS

A. Preventing Echo Chambers

Figure 3 shows the resulting progression of the Infect-Dublin network for the first experiment setup with a target neutral bias of $[0.5, 0.5, 0.5]$. The Strategic Agent (the larger vertex with a grey border) starts in the blue clique by Girvan-Neuman assignment but then attaches to yellow due to the greater numbers, changing its opinion to match. At time-step 10, the echo chamber is nearly complete and then completes by time-step 20, aside from the few blue clique edges still attached to the Strategic Agent. Green vertices show in the yellow clique as the Strategic Agent reduces

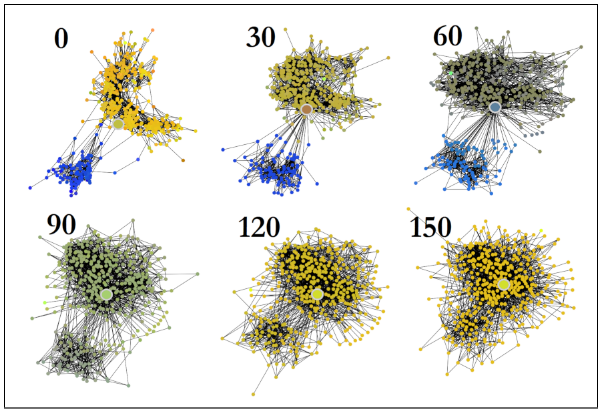

Fig. 5. The progression of the Infect-Dublin network over 150 iterations showing the Strategic Agent first driving the entire network towards the neutral bias of $[0.5, 0.5, 0.5]$ (grey) before driving it to the final target bias of $[0.9, 0.75, 0.1]$ (yellow).

x_R opinions and increases x_B opinions. This allows new edges to bridge the yellow and blue cliques by times-step 40 as opinions are less polarized and more tolerable to each other. The Strategic Agent can now change priorities to the blue clique to make their opinions less polarized. Figure 4 shows the change in the Strategic Agent's internal opinion bias over time for this experiment, showing the changes at the time-steps mentioned above. Out of twenty experiments run with the first experiment setup, the Strategic Agent successfully drives the agents towards neutral opinion and repairs connection between the cliques each time, enabling consensus of opinion.

B. Directly Driving the Network to a Target Bias

Next, we investigate if the Strategic Agent can directly and collectively drive the opinions of the agents in the network to any desired bias. Table I shows the success rates of the various target biases. The Strategic agent fails to drive the network to the first target bias, yellow's $[0.9, 0.75, 0.1]$, and succeeds in both blue's $[0.1, 0.25, 0.9]$ and green's $[0.4, 0.8, 0.4]$ over 20 experiments. Upon further experimentation, we discover that the only target opinion biases the Strategic Agent can successfully drive the network towards are those that are tolerable to the blue clique, meaning there are those in the blue clique whose opinions are no more than θ_r from the target bias. We find that the Strategic Agent begins to have success with target biases $x_R = 0.6$ and $x_B = 0.4$, or in other words, a difference of 0.5 from the template bias of the blue clique (see the last row of Table I). The difference of 0.5 is larger than θ_r, but due to the variance of initialization, there are agents who have less extreme opinions compared to the standard blue clique template.

The inability of the Strategic agent to directly drive collective opinions to some target biases is a result of using the Maximum Overlapping Interval algorithm to decide the Strategic Agent's bias each time-step. The blue clique is much smaller than the yellow clique, so the Strategic Agent always attempts to appease the yellow clique first. This

TABLE I
THE SUCCESS RATE OF THE STRATEGIC AGENT DRIVING THE ENTIRE NETWORK DIRECTLY TOWARDS A TARGET BIAS

Target Bias	Success Rate
$[0.5, 0.5, 0.5]$	20/20
$[0.9, 0.75, 0.1]$	0/20
$[0.1, 0.25, 0.9]$	20/20
$[0.4, 0.8, 0.4]$	20/20
$[0.6, 0.75, 0.4]$	3/20

means that unless the Strategic Agent is able to first drive the largest of the cliques towards a bias that is tolerable to the smallest of the cliques, it is less likely that the Strategic Agent can drive the smallest cliques to any desired opinion.

C. Two-step Target Opinion Biases

In this final experiment, the Strategic Agent first adopts a neutral target bias of $[0.5, 0.5, 0.5]$ before switching to the formerly failed target bias of $[0.9, 0.75, 0.1]$ once the cliques have collective opinions the other can tolerate. Figure 5 shows the progression of the entire network towards the previously unsuccessful yellow bias, $[0.9, 0.75, 0.1]$. Figure 6 shows the change in the Strategic Agent's internal bias over time. Similar to what we saw in Figure 4 above, until iteration 39, the Strategic Agent appeases the yellow clique to drive it's bias towards neutral. Afterwards, it shifts to influencing the blue clique. At iteration 71, the collective biases of both cliques are tolerable with each other, and so the Strategic Agent shifts to its second target bias. Following this 2-step target bias approach, the Strategic Agent succeeds in driving the entire network to the yellow bias templates in all twenty attempts.

VI. DISCUSSION

The results show that should a controller agent have a profile of the ideological stances of the people in a social media network, something that may be automatically generated by analyzing public posts made by the individuals and their publicly available account information, it can fabricate posts that can not only make it a major social media influencer, but also once its influence is large enough be able to direct public opinion. We can imagine an automated bot account that uses a language generation tool like ChatGPT to produce tweets that follow a desired ideological skew. Simply by creating posts it could generate a following of those who agree with the ideological stance without the need to directly interact with potential followers. Research by Chen et. al. shows that partisan social media accounts who subscribe to certain political ideologies are more likely to follow automated accounts that post content aligning with that subscribed bias [28], making them potentially more susceptible to such influence. Research by Hjouji et. al. shows that bots do have an effect on opinion in social networks [29]. Through a combination of the reassessment of its followers' profiles as its own skew changes and gaining new followers who can now accept the shifted ideological stance, this bot account

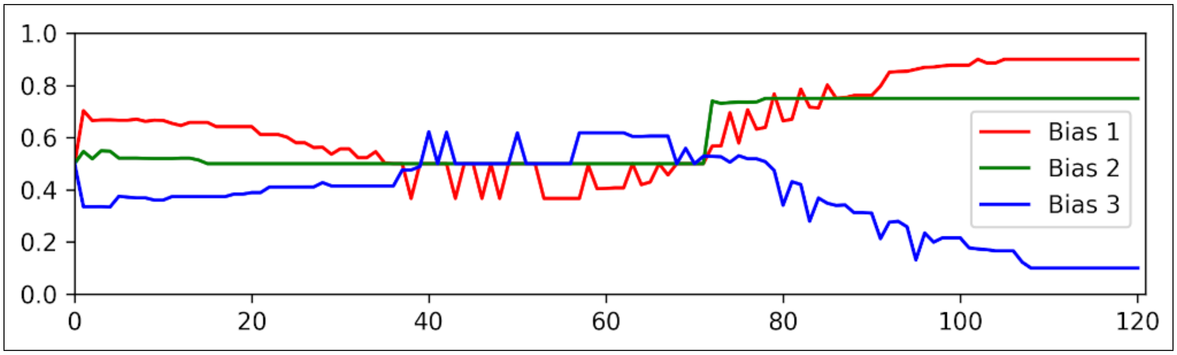

Fig. 6. Internal bias values for the Strategic Agent from the 2-step target bias experiment shown in Figure 5. At iteration 71, the collective biases of the agents in the network are tolerant enough of each other that the Strategic Agent then begins driving them to the target bias of $[0.9, 0.75, 0.1]$ (yellow).

could theoretically sway public opinion on a stance such as climate change, gun law reform, trust in vaccines, and even support for a political candidate.

Social media is a powerful platform for the sharing of information and the expression of ideas. By using social media, we are both empowering ourselves with the ability to influence others while at the same time also making ourselves susceptible to the influence of others. As the capabilities of natural language interpretation and generation technologies improve, we are providing automated systems the same power. As they improve to become indistinguishable from a human, it becomes ever more important to understand how influence over social media spreads and how to prevent those with malicious intent from using social media to manipulate public opinion.

References

[1] W. Hart, D. Albarracín, A. H. Eagly, I. Brechan, M. J. Lindberg, and L. Merrill, "Feeling validated versus being correct: a meta-analysis of selective exposure to information." *Psychological bulletin*, vol. 135, no. 4, p. 555, 2009.

[2] P. Törnberg, "Echo chambers and viral misinformation: Modeling fake news as complex contagion," *PLoS one*, vol. 13, no. 9, p. e0203958, 2018.

[3] W. E. Forum, "The global risks report 2023," *The Global Risks Report*, vol. 18, 2023.

[4] J. Scott, "Trend report social network analysis," *Sociology*, pp. 109–127, 1988.

[5] M. O. Jackson and L. Yariv, "Diffusion on social networks," *Economie Publique (Public Economics)*, vol. 16, no. 1, pp. 3–16, 2005.

[6] Y. Jiang and J. Jiang, "Diffusion in social networks: A multiagent perspective," *"IEEE Trans. Syst., Man, Cybern. A"*, vol. 45, no. 2, pp. 198–213, 2014.

[7] Y. Wang, A. V. Vasilakos, J. Ma, and N. Xiong, "On studying the impact of uncertainty on behavior diffusion in social networks," *"IEEE Trans. Syst., Man, Cybern. A"*, vol. 45, no. 2, pp. 185–197, 2014.

[8] C. Lagnier, L. Denoyer, E. Gaussier, and P. Gallinari, "Predicting information diffusion in social networks using content and user's profiles," *Advances in Information Retrieval: 35th European Conference on IR Research*, vol. 7814, pp. 74–85, 03 2013.

[9] G. D'Agostino, F. D'Antonio, A. De Nicola, and S. Tucci, "Interests diffusion in social networks," *Physica A: Statistical Mechanics and its Applications*, vol. 436, pp. 443–461, 2015.

[10] A. V. Proskurnikov and R. Tempo, "A tutorial on modeling and analysis of dynamic social networks. part 1," *Annual Reviews in Control*, vol. 43, pp. 65–79, 2017.

[11] ——, "A tutorial on modeling and analysis of dynamic social networks. part 2," *Annual Reviews in Control*, vol. 45, pp. 166–190, 2018.

[12] M. Taylor, "Towards a mathematical theory of influence and attitude change," *Human Relations*, vol. 21, no. 2, pp. 121–139, 1968.

[13] N. E. Friedkin and E. C. Johnsen, "Social influence and opinions," *Journal of Mathematical Sociology*, vol. 15, no. 3-4, pp. 193–206, 1990.

[14] R. Bredereck and E. Elkind, "Manipulating opinion diffusion in social networks," in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17*, 2017, pp. 894–900. [Online]. Available: https://doi.org/10.24963/ijcai.2017/124

[15] H. Hu, "Competing opinion diffusion on social networks," *Royal Society Open Science*, vol. 4, no. 11, p. 171160, 2017.

[16] H. Z. Brooks and M. A. Porter, "A model for the influence of media on the ideology of content in online social networks," *Physical Review Research*, vol. 2, no. 2, pp. 023 041–1–023 041–20, 2020.

[17] E. Herrera-Viedma, F. Cabrerizo, F. Chiclana, J. Wu, M. Cobo, and K. Samouylov, "Consensus in group decision making and social networks," *Studies in Informatics and Control*, vol. 26, pp. 259–268, 09 2017.

[18] J. Wu, F. Chiclana, H. Fujita, and E. Herrera-Viedma, "A visual interaction consensus model for social network group decision making with trust propagation," *Knowledge-Based Systems*, vol. 122, pp. 39–50, 2017.

[19] D. Acemoglu, G. Como, F. Fagnani, and A. Ozdaglar, "Opinion fluctuations and persistent disagreement in social networks," *Proceedings of the IEEE Conference on Decision and Control*, pp. 2347–2352, 12 2011.

[20] L. Li, A. Scaglione, A. Swami, and Q. Zhao, "Consensus, polarization and clustering of opinions in social networks," *"IEEE J. Select. Areas Commun."*, vol. 31, no. 6, pp. 1072–1083, 06 2013.

[21] X. Wang, A. D. Sirianni, S. Tang, Z. Zheng, and F. Fu, "Public discourse and social network echo chambers driven by socio-cognitive biases," *Physical Review X*, vol. 10, no. 4, p. 041042, 2020.

[22] C. Tokita and C. Tarnita, "Social influence and interaction bias can drive emergent behavioural specialization and modular social networks across systems," *Journal of The Royal Society Interface*, vol. 17, p. 20190564, 01 2020.

[23] A. V. Proskurnikov, A. S. Matveev, and M. Cao, "Opinion dynamics in social networks with hostile camps: Consensus vs. polarization," *IEEE Trans. Automat. Contr.*, vol. 61, no. 6, pp. 1524–1536, 2015.

[24] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, "Modeling echo chambers and polarization dynamics in social networks," *Physical Review Letters*, vol. 124, no. 4, p. 048301, 2020.

[25] OpenAI, "Chatgpt: Optimizing language models for dialogue," *OpenAI*, 2022. [Online]. Available: https://openai.com/blog/chatgpt/

[26] R. A. Rossi and N. K. Ahmed, "The network data repository with interactive graph analytics and visualization," in *AAAI*, 2015. [Online]. Available: https://networkrepository.com

[27] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the national academy of sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[28] W. Chen, D. Pacheco, K.-C. Yang, and F. Menczer, "Neutral bots probe political bias on social media," *Nature communications*, vol. 12, no. 1, p. 5580, 2021.

[29] Z. Hjouji, D. S. Hunter, N. Mesnards, and T. Zaman, "The impact of bots on opinions in social networks," *arXiv preprint arXiv:1810.12398*, 2018.