The Volatility of Weak Ties: Co-evolution of Selection and Influence in Social Networks

Jie Gao
Department of Computer Science, Stony Brook University
Stony Brook, NY
jgao@cs.stonybrook.edu

Grant Schoenebeck
Department of EECS, University of Michigan
Ann Arbor, MI
schoeneb@umich.edu

Fang-Yi Yu
Department of EECS, University of Michigan
Ann Arbor, MI
fayu@umich.edu

ABSTRACT

In this work we look at opinion formation and the effects of two phenomena both of which promote consensus between agents connected by ties: *influence*, agents changing their opinions to match their neighbors; and *selection*, agents re-wiring to connect to new agents when the existing neighbor has a different opinion. In our agent-based model, we assume that only weak ties can be rewired and strong ties do not change. The network structure as well as the opinion landscape thus co-evolve with two important parameters: the probability of influence versus selection; and the fraction of strong ties versus weak ties. Using empirical and theoretical methodologies we discovered that on a two-dimensional spatial network:

- With no/low selection the presence of weak ties enables fast consensus. This conforms with the classical theory that weak ties are helpful for quickly mixing and spreading information, and strong ties alone act much more slowly.
- With high selection, too many weak ties inhibit any consensus at all—the graph partitions. The weak ties reinforce the differing opinions rather than mixing them. However, sufficiently many strong ties promote convergence, though at a slower pace.

We additionally test the aforementioned results using a real network. Our study relates two theoretical ideas: the strength of weak ties—that weak ties are useful for spreading information; and the idea of echo chambers or filter bubbles, that people are typically bombarded by the opinions of like-minded individuals. The difference is in how (much) selection operates.

CCS CONCEPTS

• **Theory of computation** → **Social networks**; *Multi-agent learning*; • **Information systems** → Social networks;

KEYWORDS

Opinion formation; Polarization; Echo chambers; Tie strength; Social network structure

ACM Reference Format:

Jie Gao, Grant Schoenebeck, and Fang-Yi Yu. 2019. The Volatility of Weak Ties: Co-evolution of Selection and Influence in Social Networks. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 9 pages.

Jie Gao would like to acknowledge support by NSF DMS-1737812, CNS 1618391 and CCF 1535900. Grant Schoenebeck and Fang-Yi Yu gratefully acknowledge the support of the NSF under CAREER Award 1452915 and AitF Award 1535912.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

1 INTRODUCTION

Social ties are not static, they evolve over time and the evolution is driven by two processes. One is *selection* where an individual may seek out and form new ties; often with others that have similar attributes [31]. The other social process is *influence* in which two individuals already connected by a social tie may influence one another and converge on their personal attributes (interest, tastes, etc) [28, 32]. Both of them result in neighboring nodes being more similar than two random nodes.

The sociology literature has, for a long time, acknowledged and studied the difference of social ties [10, 11, 18]. Strong ties refer to the ties that people regularly spend effort to maintain, such as family members, close friends, and colleagues. Weak ties, on the other hand, are relatively effortless to keep and typically are much more numerous than strong ties. The difference in the type of ties is also reflected structurally. Strong ties tend to be clustered with a high clustering coefficient, while weak ties are important bridges that connect remote communities. In the seminal paper "The Strength of the Weak Ties" Granovetter [11] showed how information spreads through weak ties. While strong ties connect people who are more similar to each other (due to homophily), weak ties tend to bring fresh information to a social group, which can be extremely valuable, for example, in the case of looking for new jobs.

One of the interesting aspects of this paper is to examine the evolution of strong ties and weak ties, with selection and influence considered. By definition, strong ties and weak ties also differ in their stability or fragility. The physical constraints that form a strong tie are often stable in time and are hard to change. Many of the strong ties are not formed by selection. We are born with family ties and they stay with us for a lifetime except in extreme cases. Neighbors and colleagues are also relatively hard to change without some serious effort or cost. But weak ties, especially those discovered on a social platform, are a lot easier to form or break, making it convenient to block opinions that one does not like and stay in a comfortable "echo chamber" [2, 34].

The political science literature has confirmed the observation of geographical segregation and partisan alignment [9, 22] and of 'ideology sorting', that people tend to "segregate themselves into their own political worlds, blocking out discordant voices and surrounding themselves with reassuring news and companions" [4]. In the on-line setting, the sorting process can possibly happen at a much faster rate and a larger scale [2, 5, 14, 20, 23]. Online forums allow people to seek out like-minded individuals, including those holding unpopular views that have been shunned elsewhere [7]. Moreover, social media research clearly shows that unfriending on Facebook [30] and unfollowing on Twitter [17, 35] disproportionately affect weak ties as compared to strong ties. Between 16% and 26% of American SNS users have disconnected a tie for reasons to do with politics [14, 27, 30]. While such selection processes indeed limits the information input to certain users, it was also observed that the disconnections helped to sustain user participation in the social network [20].

Our Approach. In this work we develop a model of opinion formation and changes with two competing opinions/behaviors. Examples include political views (liberal v.s. conservative) or behaviors (smoking/non-smoking, drug use/no drug use). The opinions are influenced by one's friends which could be connected by strong ties or weak ties. Generally speaking, one's opinion is going to move toward the majority opinion in his/her friend circle over time. Meanwhile, selection may also happen such that a node re-wires ties when he/she has different opinion from his/her friends. In our model, we assume that only the weak ties can be rewired and strong ties do not change. The network structure as well as the opinion landscape thus co-evolve with two important parameters: p_{select}, the probability of a selection as the next action as opposed to influence; and q_{strong}, the fraction of strong ties in the network. The objective of this paper is to answer the following question: does the opinion distribution converge and if so how fast does it converge with respect to the two parameters?

Related Work. There has been work on co-evolution of social ties and opinions without separating strong/weak ties. Holme and Newman [13]

show a phase transition from a segregated network to a homogeneous network, controlled by p_{select}. Durrett et al. [8] built on top of the Holme-Newman model and considered two models of selection: rewire-to-random, and rewire-to-same. Cohen et al. [6] study a problem of opinion formation with continuous values with influence and selection. Kempe et al. [16] considered agents with multiple dimensions/attribute types and only agents who are similar in many dimensions can influence each other. They characterized the equilibrium outcome and proved convergence.

An expansive literature attempted to validate selection and influence models using real-world data, although some of them are limited as they assume independent observations and no external factors [31]. Lewis et al. [19] considered Facebook data and discovered that there could be a large variation of whether selection or influence is more prominent, depending on the studied attributes. Further, selection and influence can be heavily entangled. For example, in a static network (when selection does not exist), both cooperative and selfish behaviors are contagious. But in a dynamic network, selfish behavior is still contagious, but cooperative behavior is not [15]. Thus selection and influence in network co-evolution definitely deserve further study in different social settings.

2 MODEL

2.1 Model of Agent Network

To encode the interaction among people, we use a directed graph $G = (V, E_S, E_W)$ with V as the set of nodes and two types of edges—*strong ties*, E_S, and *weak ties*, E_W. For $v \in V$ let $d_S(v)$ be the *out-degree* of strong ties of node v and the i-th strong out-neighbor of node v is denoted by $\delta_S(v)_i$. We define $d_W(v)$ and $\delta_W(v)_i$ analogously. We allow multi-edges and self loops in both E_S and E_W.

2.2 Dynamics of Influence and Selection

Each agent $v \in V$ has an opinion $\chi(v) \in \{0, 1\}$. We call $\chi = \{\chi(v) : v \in V\}$ the *opinion vector*. For $\sigma \in \{0, 1\}$, let $\boldsymbol{x}(\sigma) \triangleq \{v \in V : \chi(v) = \sigma\} \subseteq V$ denote the set of nodes with opinion σ. Let $R_S^{\chi}(v) = \frac{|\{i:\chi(\delta_S(v)_i)=1\}|}{d_S(v)}$ be the fraction of strong ties which have an endpoint with opinion 1, and similarly define $R_W^{\chi}(v)$.

The process Sel-Inf$(G^{(0)}, f_{\text{inf}}, p_{\text{select}}, q_{\text{strong}})$ is a discrete time Markov chain over state space $\{(\chi, G)\}$ where $G^{(0)}$ is the initial network of agents, $f_{\text{inf}} : [0, 1] \mapsto [0, 1]$ is an influence function, parameter $p_{\text{select}} \in [0, 1)$ denotes the amount of selection (versus influence), and $q_{\text{strong}} \in [0, 1]$ denotes the influence of the strong ties (versus weak ties). To this end we define $R^{\chi}(v) = q_{\text{strong}} R_S^{\chi}(v) + (1 - q_{\text{strong}}) R_W^{\chi}(v)$ to be the weighted fraction of v's neighbors that are 1.

The dynamics Sel-Inf$(G^{(0)}, f_{\text{inf}}, p_{\text{select}}, q_{\text{strong}})$ start with the graph $G^{(0)}$ and initial opinions that are uniformly and independently randomly selected.

Given state $Y^{(t)} = (\chi^{(t)}, G^{(t)})$ at time t, the dynamics updates to $Y^{(t+1)}$ as follows: initially set $Y^{(t+1)} = Y^{(t)}$, choose an agent v uniformly at random and update $Y^{(t+1)}$ with one of the following two operations:

Selection. With probability p_{select}, agent v randomly chooses a weak tie and rewires if they disagree: select a random number k between $1, \ldots, d_W(v)$, and let $u = \delta_W^{(t)}(v)_k$. Then

$$\delta_W^{(t+1)}(v)_k = \begin{cases} u & \text{, if } \chi^{(t)}(v) = \chi^{(t)}(u) \\ \text{a random node in } V & \text{, otherwise.} \end{cases} \tag{1}$$

Influence.[29] Otherwise (with probability $1 - p_{\text{select}}$), agent u updates its opinion,

$$\chi_v^{(t+1)} = \begin{cases} 1 & \text{with probability } f_{\text{inf}}\left(R^{\chi^{(t)}}(v)\right) \\ 0 & \text{, otherwise.} \end{cases} \tag{2}$$

recall that $R^{\chi^{(t)}}(v)$ is the q_{strong} weighted fraction of v's neighbors with opinion 1 at time t.

We say the process reaches *consensus* if all agents have the same opinion, and we use the number of influence steps as the consensus time.

Remark 2.1. Our model is similar to the Holme-Newman model [13]. In the selection phase of our model, the chosen node picks a random edge, and when the endpoint has a different opinion rewires the edge to a random node (rewired when disagreeing). In their model, a random edge is rewired to a random node with the same opinion (rewired to the same). For the influence phase, their model uses the voter model to update opinions.

Remark 2.2. We will describe our simulation results using ρ_{select} instead of p_{select} where $p_{\text{select}} = \frac{d\rho_{\text{select}}}{1+(d-1)\rho_{\text{select}}}$ and d is the average degree of the graph. Here ρ_{select} just rescales p_{select} to correctly normalize for the degree. This way, if v is a node of degree d, the rate that the opinion of $\delta_W(v)_i$ is updated via selection versus influence is ρ_{select} versus $1 - \rho_{\text{select}}$ and does not depend on d.

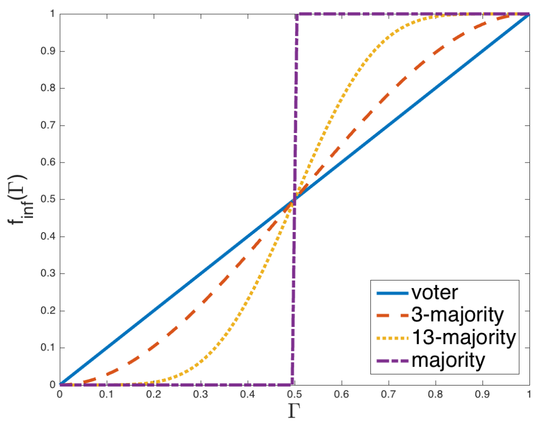


Figure 1: The function f_{inf} for different influence dynamics. The k-majority model, with an increasing k, changes from the voter model to the majority model.

2.3 Choices of Influence dynamics

We consider k-majority dynamics (choose k neighbors according to their edge weights independently with replacement and change the opinion to the majority opinion of these k neighbors),

$$f_{\text{inf}}(x) = \sum_{\ell=\lceil k/2 \rceil}^{k} \binom{k}{\ell} x^{\ell}(1-x)^{k-\ell}. \tag{3}$$

This generalizes several previously studied models:

- Voter Model ($k = 1$): agent u chooses a neighbor v with probability proportional to the weight and updates to v's opinion, $f_{\text{inf}}(x) = x$ [12].
- Majority ($k \to \infty$): agent u updates to the opinion with maximum weight, when there is a tie, the opinion is chosen at random [24].
- 3-majority dynamics ($k = 3$): agent u polls the opinion from three random neighbors and takes the majority as the new opinion [3].

For $k > 1$ this family of influence dynamics can be seen as the smooth version of majority dynamic with "the rich get richer property"— if $R_u > 1/2$, more than half of u's neighbors are 1 then the probability that agent u updates to 1 is greater than R_u, the fraction of u's 1 neighbors; moreover on a complete graph if the number of agents with opinion 1 is greater than the number of agents with opinion 0 there is a "drift" for opinion 1 such that the number of agents with opinion 1 tends to increase. We are primarily interested in the case where $k > 1$, but include the $k = 1$ case for contrast.

2.4 Our Problem

In this paper we try to understand the role of weak ties in promoting consensus with two main parameters: ρ_{select}, the probability of selection as the next action as opposed to influence; and q_{strong}, the fractional influence of the strong ties in the network. We consider the entire parameter space: $\rho_{\text{select}} \in [0, 1)$ and $q_{\text{strong}} \in [0, 1]$. For shorthand, we refer to this as Sel-Inf($f_{\text{inf}}, p_{\text{select}}, q_{\text{strong}}$), when the graph is clear.

In this paper we consider a number of graph topologies, networks generated by the Newman-Watts model and a real-world ego-network from Facebook [21].

3 SPATIAL NETWORKS

3.1 Simulation setting

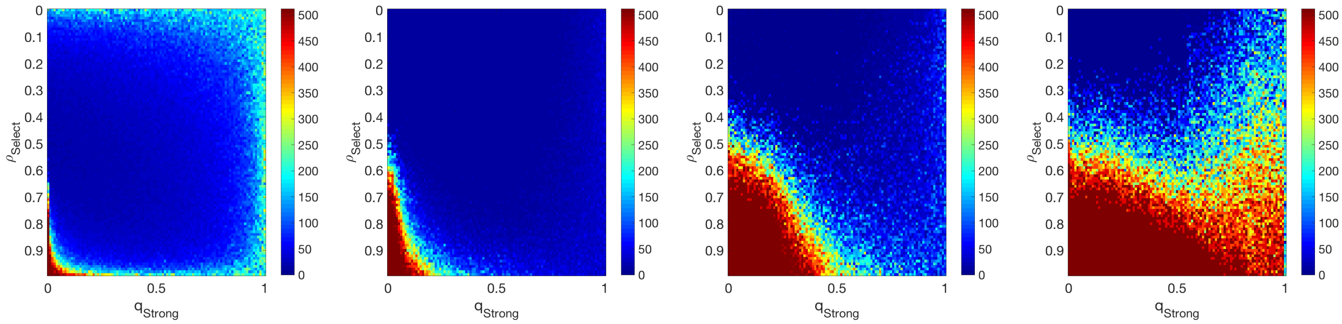
In this section, the initial graph we study is based on the Newman-Watts model [25]. The nodes form a two dimensional lattice wrapped into a torus. Each node has 12 strong ties connecting it to nodes with Hamming distance less than 2, and 10 weak ties to random nodes drawn uniformly and independently with replacement.

We run simulations on networks of size ranging from 16×16 to 64×64 (256 to 4096 nodes). A representative figure on the number of influence steps until consensus is shown in Figure 2. The color at each point ($\rho_{\text{select}}, q_{\text{strong}}$) represents the number of influence steps before consensus (or timeout) normalized by the the size of the graph and averaged among the trials of the dynamics Sel-Inf($f_{\text{inf}}, p_{\text{select}}, q_{\text{strong}}$). We stop the dynamics if the total number of influence steps is more than twice the square of the size of the graph. In the larger graph, this corresponds to 33,554,432 influence steps and, for some parameter settings, over 10 billion total steps. For the 256 node graph, we run 10 trials for each of 100×101 parameter settings. For the 4096 node graph, we run 5 trials for each of 50×51 parameter settings.

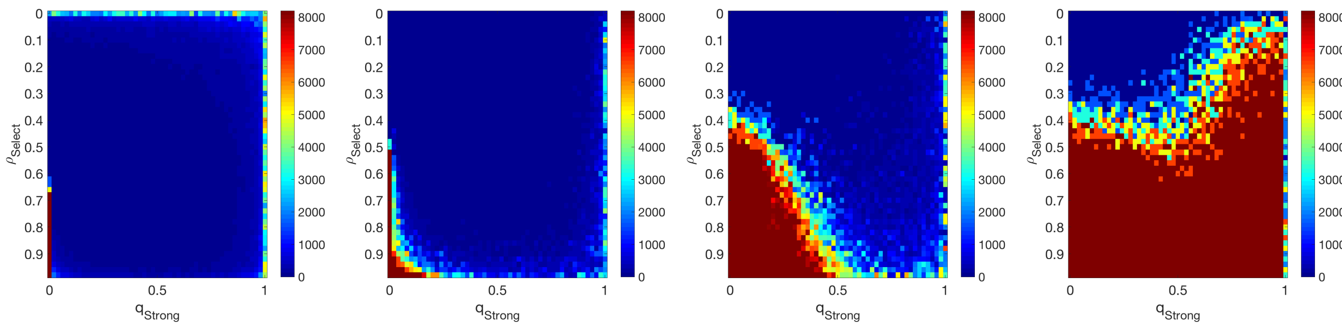
3.2 Simulation Results Overview

To better understand Figure 2, we first consider what happens with different selection rates. When $\rho_{\text{select}} < 0.5$, which is the upper part of the plots, the majority-like processes (3-majority, 13-majority, and majority) reach consensus faster if the weight of weak ties is larger (q_{strong} being smaller). This is natural because the graph topology is more stable when ρ_{select} is small. Once the number of nodes with different opinions become imbalanced *the weak ties act like sampling a complete graph and help the opinions to mix, strengthening the imbalance.* If q_{strong} is close to 1, the network has mostly only the strong ties that connect local neighbors. Even though there may exist a global imbalance of opinions, it still takes a long time to spread this imbalance through strong ties.

However, when selection rate is high ($\rho_{\text{select}} > 0.5$, the lower part of the plots), the majority-like processes (3-majority, 13-majority, and majority) reach consensus slower or even get stuck if there are a large fraction of weak ties (when q_{strong} is small). In contrast to the low selection setting, here the weak tie weights are frequently updated and form stronger connections among the agents with the same opinion. Informally, *the weak ties form community structures*



(a) Voter Model (k=1) (b) 3-Majority (c) 13-Majority (d) Majority(16×16)



(a) Voter Model (k=1) (b) 3-Majority (c) 13-Majority (d) Majority(64×64)

Figure 2: Consensus time on spatial network. The color at each point $(\rho_{\text{select}}, q_{\text{strong}})$ in this bit map represents the average number of influence steps before consensus (or timeout). The size of graph in the top row is 256 and the bottom row is 4096.

which hinder the agents from communicating between different opinions and prevent the opinions from mixing. As a result, the higher the selection rate is, the harder for the agents to reach consensus.

We hypothesize that there are three distinct theoretical cases:

Fast Consensus Consensus takes a logarithmic number of steps (per node).

Slow Consensus Consensus is reached in polynomial time.

No Consensus Consensus is either never reached or takes exponential time.

Roughly speaking: we expect fast consensus is represented by the deep blue region; no consensus by the deep red region; and slow consensus by the other colors. Notice that when there are no strong ties ($q_{\text{strong}} = 0$) the transition from fast consensus to no consensus is rapid. We hypothesize that the there is a threshold here. Moreover, that there is a "triple point" incident on each of these three regions.

In the remainder of our analysis we focus on the three "edges": either $q_{\text{strong}} = 0$ or $\rho_{\text{select}} \in \{0, 1\}$, and we change the other parameters. Note that when $q_{\text{strong}} = 1$ selection cannot operate and the value of $\rho_{\text{select}} \in \{0, 1\}$ is immaterial. So this case is omitted.

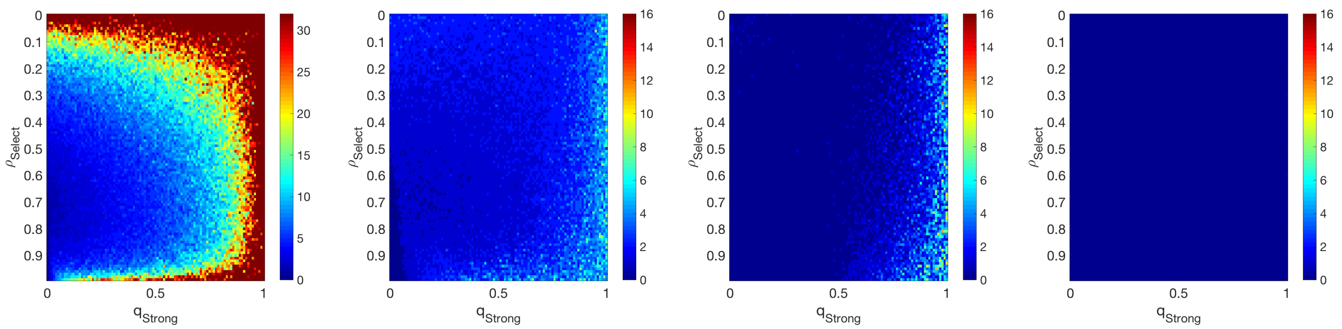
3.3 Weak Ties Only ($q_{\text{strong}} = 0$)

In this section we study the effects of the relative frequency between selection and influence (ρ_{select}) on the consensus time of Sel-Inf$(f_{\text{inf}}, p_{\text{select}}, q_{\text{strong}})$ when the strong ties are absent, $q_{\text{strong}} = 0$. This corresponds to the left edge of the plots in Figure 2.

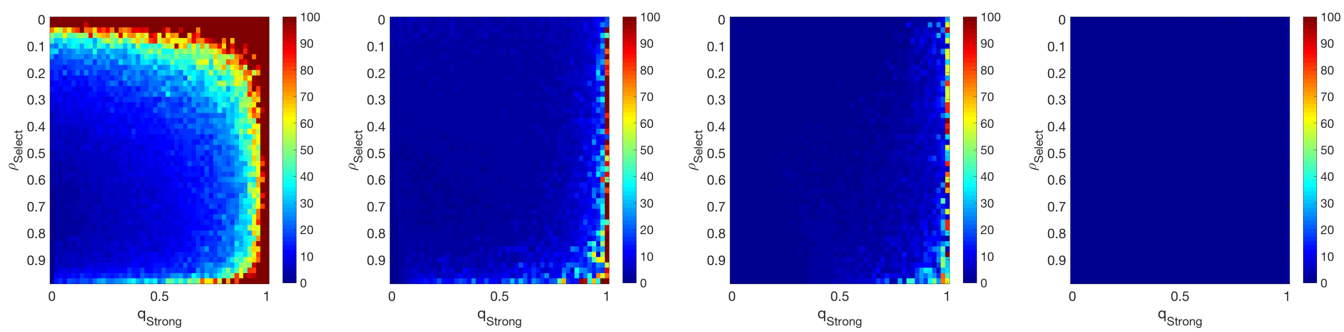
We can see that if $\rho_{\text{select}} = 0$, then the dynamics quickly converge in all but the voter model, where it slowly converges. On the other hand if $\rho_{\text{select}} \to 1$, then it nearly always times out before converging. We hypothesize that in this case there is no consensus. One way we can see this is in Figure 3, which plots the number of times nodes switch opinions, normalized by the size of the graph, before the processes reach consensus. A switch is an influence step when the chosen agent changes its opinion. The total number of switches is quite small in this region. This indicates that no real progress is being made.

k > 1. First we consider $k > 1$—recall f_{inf} is k-majority. We see that on the left side of the plots in Figure 2 the time quickly transitions from fast to very slow. Again the data in Figure 3 backs up the story that the process transitions from making quick progress (with few switches) to making no progress (with a lot of switches). In the following section we use theoretical analysis to show that in the mean field approximation the k-majority dynamics (for odd k) converges to *segregation* if the relative frequency of selection is high enough. We present theoretical results on the mean field approximation of this setting in Section 4.

k = 1. Turning toward the case $k = 1$, we notice a large difference. Here the dynamics appear to converge slowly at $\rho_{\text{select}} = 0$. The time to consensus is intermediate (Figure 2), and requires many switches (Figure 3). However, as ρ_{select} increases, the process transitions to fast consensus (fast time and few switches). Finally, as



(a) Voter Model (b) 3-Majority (c) 13-Majority (d) Majority



(a) Voter Model (k=1) (b) 3-Majority (c) 13-Majority (d) Majority

Figure 3: Switches on Spacial Network. The color at each point ($\rho_{\text{select}}, q_{\text{strong}}$) in this a bit map represents the total number of switches (before consensus or timing out) normalized by the size of the network for Sel-Inf($f_{\text{inf}}, p_{\text{select}}, q_{\text{strong}}$)**. The size of graph in the top row is** 256 **and the bottom row is** 4096**.**

ρ_{select} continues to increase we transition to increasingly timing out (slow time and few switches). The slow consensus at $\rho_{\text{select}} = 0$ is expected, because the voter model has no drift. However, the fast consensus time for intermediate values of ρ_{select} is surprising. We hypothesize that it is due to the details of the selection process which induces a rich-get-richer drift. When updating, if a node is in the minority, then its selections acts slower (because the updates are additive, but the total mass of its weak ties is smaller). This means that minority nodes are more likely to be connected to majority nodes than vice versa.

3.4 No Selection, Only Influence ($\rho_{\text{select}} = 0$)

In this section, we consider the setting when there is no selection. Therefore the process boils down to influence in a static network with strong and weak ties. The results are at the top edge of the plots in Figure 2.

For k-majority models for $k > 1$, we hypothesize that any non-zero fraction of weak ties leads to fast consensus, which is supported in the simulation results. The reason is that as soon as an opinion is a global leader, the weak ties introduce a global drift. Since there is no selection, each node connects uniformly to all nodes via weak ties. The strong ties can make local imbalances, but these cancel each other out as the size of the "boundary" for each opinion is necessarily the same. In Figure 3, the number of switches increases when there are more strong ties (with q_{strong} increasing). When q_{strong} is small, on average each node switches fewer than 4 times before consensus is reached — weak ties help to spread the imbalance of opinion quickly and in most of the influence steps the chosen agent updates to the global majority correctly.

However, with just strong ties ($q_{\text{strong}} = 1$, the top right corner), the process predominantly changes only at the boundary of regions of different opinions. Since the boundary of each opinion is the same, the process takes an unbiased walk (without drift) and converges slowly.

For $k = 1$, we have the voter model, which has no drift regardless of q_{strong}. However, as there are more weak ties, the graph mixes better and convergence speeds increase slightly. Indeed, as the fraction of strong ties increases, the number of switches in Figure 3 increases. However, compared to majority-like dynamics the voter model has a much larger number of switches regardless of the value of q_{strong}.

3.5 Lots of Selection ($\rho_{\text{select}} \to 1$)

In this section, we want to understand when ρ_{select} is nearly 1, which is near the bottom edge of the plots in Figure 2. When $\rho_{\text{select}} = 1$, i.e., no influence, the opinions do not change. Thus the network does not reach consensus.

When q_{strong} and ρ_{select} is nearly 1 (near the right bottom corner), there are no weak ties. Although almost all actions are selections, there are simply no weak ties to work on, and so the selection steps do not affect. (Note that Figure 2 only counts influence steps.) Thus, as discussed in the earlier section it converges but slowly.

When $\rho_{\text{select}} \to 1$ and q_{strong} is increasing, the strong ties increasingly help with consensus, but the weak ties are almost surely connecting nodes of the same opinion. Conversely, as the number of weak ties increases, they increasingly promote segregation.

For the majority model, it is abruptly not stuck when $q_{\text{strong}} = 1$. Here it is, in theory, possible that the dynamics get stuck (for example if an 8×16 region of nodes in the torus have opinion 0 and the other 8×16 region have opinion 1. All agents will have three neighbors of their type. However, in our empirical results, these trials never do become stuck. Since there are only strong ties, we hypothesize, that in the case the dynamics do converge it cannot be done quickly (in logarithmic time per node) but must take a polynomial time per node to converge.

4 THEORETICAL RESULTS

In this section, we analyze the process Sel-Inf when the d-regular random graph which only has weak ties, and we show the mean field approximation process converges to segregation when the selection rate is higher than a certain threshold which depends on the influence function f_{inf} and the degree d.

Formally, we consider Sel-Inf$(G^{(0)}, f_{\text{inf}}, p_{\text{select}}, q_{\text{strong}})$ where the initial weak graph $E_W^{(0)}$ is a directed d-regular random graph (i.e., each node has d out neighbors selected at random), $q_{\text{strong}} = 0$, and f_{inf} is the k-majority influence dynamics with $k \geq 3$. We note that the nodes with the same initial state will have the same *expected* behavior. Specifically we can partition the nodes by their initial opinions into $U_0 \triangleq \boldsymbol{x}^{(0)}(0)$ and $U_1 \triangleq \boldsymbol{x}^{(0)}(1)$ and can assume $|U_0| = |U_1| = n/2$.

For $\sigma \in \{0, 1\}$ we call $v \in \boldsymbol{x}^{(0)}(\sigma)$ a *type* σ node, and similarly define type $\tau \in \{0, 1\}$ nodes. We set $X_\sigma(t)$ to be the average probability of type σ nodes having opinion 1 at time t, and $C_{\sigma,\tau}(t)$, the expected cut of the weak ties between a type σ node and a type τ node at time t. Formally,

$$\begin{cases} X_\sigma(t) \triangleq \frac{1}{|U_\sigma|} \sum_{v \in U_\sigma} \mathbb{E}\left[x_v^{(t)}\right] \\ C_{\sigma,\tau}(t) \triangleq \frac{1}{|U_\sigma|} \sum_{v \in U_\sigma} \frac{1}{d_W(v)} \mathbb{E}\left[\left|\{i : \delta_W^{(t)}(v)_i \in U_\tau\}\right|\right] \end{cases} \tag{4}$$

Theorem 4.1. *Given constants $k > 1$ odd and d, let $G^{(0)}$ be a directed d-regular random graph with n nodes, and $q_{\text{strong}} = 0$, there exists $p_{\text{select}}{}^* \in (0, 1)$ such that for all $p_{\text{select}} > p_{\text{select}}{}^*$ for sufficiently large n, the mean field approximation of* Sel-Inf$(G^{(0)}, f_{\text{inf}}, p_{\text{select}}, q_{\text{strong}})$ *defined in Equation* (4)*, the system converges to segregation:*

$$\lim_{t\to\infty} X_0(t) = 0, \lim_{t\to\infty} X_1(t) = 1 \tag{5}$$

$$\lim_{t\to\infty} C_{0,1}(t) = \lim_{t\to\infty} C_{1,0}(t) = 0. \tag{6}$$

Intuitively, this theorem shows in the mean field approximation, the cut between two sets $\boldsymbol{x}^{(0)}(0)$ and $\boldsymbol{x}^{(0)}(1)$ converges to zero, the agents in $\boldsymbol{x}^{(0)}(0)$ converge to opinion 0, and the agents in $\boldsymbol{x}^{(0)}(1)$ converge to opinion 1.

Now we give some intuitions of the proof. We first show that as n increases the recurrence relation can be (rigorously) quantitatively approximated by a system of ordinary differential equation (ODE) (c.f. Figure 4). We analyze the corresponding system of ODE using tools from dynamical systems theory. One major challenge of Theorem 4.1 is to argue the limits of system (4) converges to $(0, 0)$ without knowing their analytic solutions. We achieve this by using tools in the qualitative analysis of dynamical systems which is of independent interest.

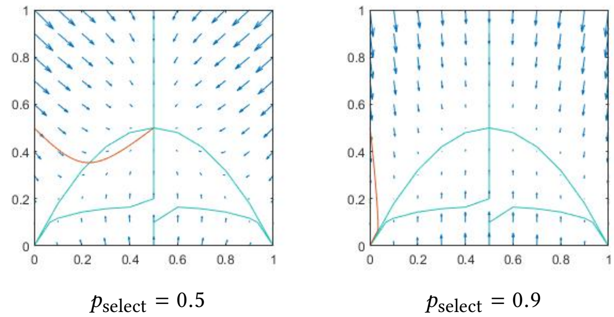


Figure 4: The vector field for dynamical system of (4) with initial condition $(X_0(t), C_{0,1}(0)) = (0, 0.5)$ for 3-majority under different p_{select}. The green lines represent the zeros of the system of differential equations, and the red path is the numerical solution of the dynamical system. On the left-hand side (small p_{select}), the dynamical system mixes and the probability of having opinion 1 and the connection between two types of nodes converges to $(0.5, 0.5)$. On the right-hand side (large p_{select}), the system segregates— the connection/cut between two types of nodes converges from 0.5 to 0 which is characterized in the Theorem 4.1.

4.1 Preliminaries

For our theoretical result, we need to introduce several results in dynamical systems. Given $f = (f_1, \ldots, f_d) : D \subseteq \mathbb{R}^d \mapsto \mathbb{R}^d$ where D is an open set with $z(0) = Z^{(0)}$, we define the following two processes:

$$\frac{d}{dt} z = f(z), \text{ and } Z^{(k+1)} - Z^{(k)} = \frac{1}{n} f\left(Z^{(k)}\right). \tag{7}$$

Theorem 4.2 (Convergence of Eular forward method [1]). *Let $f : D \to \mathbb{R}^d \in C^1$ such that the derivative f' exists and is continuous with $\|f(x)\| \leq M$, and $\|f(x,t) - f(z,t)\| \leq L\|x - z\|$. Then in Equation (7), for all $t > 0$ the $Z^{(t)}$ differs from the true solution z by at most*

$$\|Z^{(nt)} - z(t)\| \leq \frac{M}{n}(e^{Lt} - 1).$$

To capture the long term behaviour of (7), we introduce a notion of stability: an equilibrium $z^* \in D$ is *asymptotically stable* if there exists $\delta > 0$ such that $\|z(0) - z^*\| \leq \delta \Rightarrow \lim_{t\to\infty} \|z(t) - z^*\| = 0$. The stability of the system can be determined by the linearization of the system which is stated below.

Theorem 4.3 (Lyapunov's indirect method [33]). *Let $D \subset \mathbb{R}^d$, $z^* \in \mathbb{R}^d$, $f : D \mapsto \mathbb{R}^d$, and $A \in \mathbb{R}^{d\times d}$ where D is a neighborhood of the z^*, f is a continuously differentiable function C^1, z^* is an equilibrium point such that $f(z^*) = 0$, and $A = \frac{\partial f}{\partial z}|_{z=z^*}$ is the derivative of f at z^*. Then z^* is asymptotically stable if A is* Hurwitz, *so that all eigenvalue of A, λ, the real part of λ is negative, $\Re(\lambda) < 0$.*

Moreover, there exists a closed set $N \subseteq D$ and $z^* \in N$ and a potential function $V : N \to \mathbb{R}$ such that $V(z^*) = 0$, and $V(z) > 0$, $\frac{d}{dt}(V(z)) < 0$ for $z \in N \setminus z^*$.

However the above theorem only captures the behaviour of z when it is close enough to the stable fixed point z^*. To show our process converges to a neighborhood of the stable fixed point, Theorems 4.4 is useful as long as the system (7) is in the plane.

To state the theorem we need to introduce more terminologies: A set is *bounded* if it is contained in some sphere $\{z \in \mathbb{R}^d : \|z - \alpha\| < C\}$ for some $\alpha \in \mathbb{R}^d$ and $C > 0$. A point $p \in \mathbb{R}^d$ is called an *ω-limit point* of the trajectory $\gamma_{z_0} = \{z(t) : t \geq 0, z(0) = z_0\}$ of the system (7) if there is a sequence $t_n \to \infty$ such that $\lim_{n\to\infty} z(t_n) = p$.

Theorem 4.4 (Poincare-Bendixon Theorem [33]). *Let $\frac{d}{dt}z = f(z)$ be a system of differential equations defined on D an open subset in $\mathbb{R}^2$ where f is differentiable. Suppose a forward orbit with initial condition z_0, $\gamma_{z_0} = \{z(t) : t \geq 0, z(0) = z_0\}$ is bounded. Then $\omega(z_0)$ either contains a fixed point or is a periodic orbit.*

The following theorem gives us a sufficient condition for the nonexistence of a periodic orbit. Note that the theorem only holds for two dimensions systems.

Theorem 4.5 (Bendixson's Criteria [33]). *Let f be differentiable in D where D is a simply connected region in $\mathbb{R}^2$. If the divergence of the vector field f is not identically zero and does not change sign in D then $\frac{d}{dt}z = f(z)$ has no closed periodic orbit in D.*

4.2 Symmetry in Equation (4)

Note that by the definition $C_{0,0}(t) + C_{0,1}(t) = C_{1,0}(t) + C_{1,1}(t) = 1$. For all $\sigma \in \{0, 1\}$, denote the difference of a sequence (a_t) as $\Delta(a_t) \triangleq a_{t+1} - a_t$

$$\begin{aligned}\Delta(X_\sigma(t)) =& \frac{1 - p_{\text{select}}}{2|U_\sigma|}\left(f_{\inf}(R_\sigma(t)) - X_\sigma\right)\\ \Delta(C_{\sigma,\sigma'}(t)) =& \frac{p_{\text{select}}}{4d|U_\sigma|}\big[C_{\sigma,\sigma}(2X_\sigma(1 - X_\sigma))\\ & - C_{\sigma,\sigma'}(X_\sigma + X_{\sigma'} - 2X_\sigma X_{\sigma'})\big]\end{aligned}$$

where $R_\sigma(t) \triangleq C_{\sigma,\sigma}(t)X_\sigma(t) + C_{\sigma,\sigma'}(t)X_{\sigma'}(t)$ and σ' is the complement of σ such that $\sigma, \sigma' \in \{0, 1\}$ and $\sigma' \neq \sigma$.

For the initial conditions, by definition, $X_0(0) = 0$, $X_1(0) = 1$, and the initial weak graph $E_W^{(0)}$ is a directed d-regular random graph, so $C_{00}(0) = C_{01}(0) = C_{10}(0) = C_{11}(0) = 0.5$. Thus, for all $t \geq 0$, $X_0(t) = 1 - X_1(t)$, $C_{0,0}(t) = C_{1,1}(t)$, and $C_{0,1}(t) = C_{1,0}(t)$.

With these symmetries, we further define $Z^{(t)} = \left(Z_1^{(t)}, Z_2^{(t)}\right)$ where $Z_1^{(t)} \triangleq X_0(t)$ and $Z_2^{(t)} \triangleq C_{0,1}(t)$. We can reduce the number of parameters from 6 to 2 and have

$$\begin{cases} Z_1^{(t+1)} - Z_1^{(t)} = \frac{1}{n}(1 - p_{\text{select}})f_1\left(Z^{(t)}\right)\\ Z_2^{(t+1)} - Z_2^{(t)} = \frac{1}{n}\frac{p_{\text{select}}}{2d}f_2\left(Z^{(t)}\right)\end{cases} \tag{8}$$

where

$$\begin{cases} f_1(Z) = \left(f_{\inf}\left(Z_1 + Z_2(1 - 2Z_1)\right) - Z_1\right)\\ f_2(Z) = \left(-Z_2 + 2Z_1(1 - Z_1)\right)\end{cases}. \tag{9}$$

Observe that as n increases, the above process can be approximated by the following ODE by Theorem 4.2:

$$\begin{cases} \frac{d}{dt}z_1 = (1 - p_{\text{select}})f_1(z)\\ \frac{d}{dt}z_2 = \frac{p_{\text{select}}}{2d}f_2(z)\end{cases} \tag{10}$$

4.3 Proof of Theorem 4.1

The main idea of the proof has three parts:

(1) There exists a ${p_{\text{select}}}^*$ such that for all $p_{\text{select}} > {p_{\text{select}}}^*$, $Z^{(t)}$ converges to $(0, 0)$ if there is t_0 such that $Z^{(t_0)}$ is close to $(0, 0)$.
(2) Given $p_{\text{select}} > {p_{\text{select}}}^*$ there exists t_0 large enough such that z hits an asymptotically stable region for $(0, 0)$ at time t_0.
(3) Given t_0, there exists a n large enough such that $Z^{(nt_0)}$ and $z(t_0)$ are close.

We formalize these three statements in Lemmas 4.6, 4.7 and 4.9. The proof of Theorem 4.1 is deferred to the full version.

Lemma 4.6. *For all p_{select}, there exist $\delta_{p_{\text{select}}} > 0$ and large enough n such that if there is $t_0 \geq 0$, $\left\|Z^{(t_0)} - 0\right\| \leq \delta_{p_{\text{select}}}$, then*

$$\lim_{t\to\infty}\left\|Z^{(t)} - 0\right\| = 0.$$

The detailed proof is deferred to the full version. To prove Lemma 4.6, there are two parts: by Theorem 4.3, we can show 0 is asymptotically stable for (10) and there is a potential function V. Then we can show the $\mathbf{Z}^{(t)}$ in (8) converges to 0 when $\mathbf{Z}^{(0)}$ is close to 0 by showing $V\left(\mathbf{Z}^{(t)}\right)$ is decreasing as t increases when n sufficiently large.

Lemma 4.7. *There exists ${p_{\text{select}}}^* < 1$ large enough such that for all $p_{\text{select}} > {p_{\text{select}}}^*$ and $\delta > 0$, there is t_0, $\|z(t_0) - 0\| \leq \delta/3$.*

The proof of Lemma 4.7 is more complicated. The statement basically says starting from the initial condition $(0, 0.5)$, z converges to 0 when p_{select} is large enough.

Lemma 4.8 (stability). *There exists ${p_{\text{select}}}^* < 1$, a region $R_A \subset \mathbb{R}^2$ containing $(0, 0)$, and $t_0 > 0$. If $p_{\text{select}} \geq {p_{\text{select}}}^*$ and $z(0) = (0, 0.5)$, $z(t_0) \in R_A$, and $z(t) \in R_A$ for all $t \geq t_0$.*

The detailed proof is deferred to the full version. Informally, to prove the second part of Lemma 4.8, we first define our stable region $R_A = \{(x_1, x_2) : 0 \leq x_1 \leq x_1^*, 0 \leq x_2 \leq x_2^*\}$.[1] where $\left(x_1^*, x_2^*\right)$ is the fixed point of Equation (9) with smallest positive x_1^*. We must show at each boundary the drift is inward such that if the $z(t)$ is at the boundary the $z(t + \epsilon)$ will go back to the stable region. For the first part, we show z hits the stable region R_A fast by taking ${p_{\text{select}}}^*$ large enough. With Lemma 4.8 the rest of the proof of Lemma 4.7 goes as follows:

Proof of Lemma 4.7. Our system is two dimensional, so the solution z is a Jordan curve, and it is bounded in R_A if $z \in R_A$ for $t > \tau_0$ by Lemma 4.8. Therefore by Theorem 4.4 z converges to either a fixed point or a limit cycle.

[1] Technically, we need our regions to avoid the fixed point, so $R_A = [0, y_1^*] \times [0, y_2^*]$ where $y_1^* < x_1^*$ and $y_2^* < x_2^*$. By the continuity of the system and because the fixed point (x_1^*, x_2^*) is a saddle point, the stability argument still holds.

We first show no limit cycle. By Theorem 4.5, it is sufficient to show the divergence of f is not identically zero and does not change sign in R_A

$$\nabla f = (1 - p_{\text{select}}) \left(-1 + f'_{\text{inf}} (Z_1 + Z_2(1 - 2Z_1))\right) - \frac{p_{\text{select}}}{2d}.$$

Because a k-majority function defined in (3) is Lipschitz such that there exists $L_k > 0$ for all $x \in [0, 1], |f'_{\text{inf}}(x)| \leq L_k$, we can take $p_{\text{select}}{}^*$ large enough such that for all x and $p_{\text{select}} \geq p_{\text{select}}{}^*$, $\nabla f(\boldsymbol{x}) \leq (1 - p_{\text{select}})(-1 + L_k) - \frac{p_{\text{select}}}{2d} \leq (1 - p_{\text{select}})(L_k - 1 + 1/2d) - 1/2d < 0$. Since 0 is the only fixed point in R_A and there is no limit cycle, $\lim_{t \to 0} \boldsymbol{z}(t) = 0$ □

Lemma 4.9. *Given constants $t_0 \geq 0$, $\delta > 0$, and p_{select} there exists n large enough such that $\|Z^{(n\, t_0)} - \boldsymbol{z}(t_0)\| \leq \delta/3$.*

Since a k-majority function (3) is smooth, Lemma 4.9 is a corollary of Theorem 4.2.

5 REAL SOCIAL NETWORK

5.1 Simulation Setting

We use a dataset consisting of social circles (egocentric networks) collected from Facebook [21]. The graph has 4039 nodes and 88,234 edges. In this section, we only consider the 10-core[2] of Facebook graph as our base network, which contains 2987 nodes and 83,181 edges. We take V as the set of vertices of the 10-core of Facebook graph. Then we use Jaccard similarity[3] to measure tie strength and take the top 80% edges with the highest Jaccard similarity as strong ties edges, and rest as the initial weak ties.

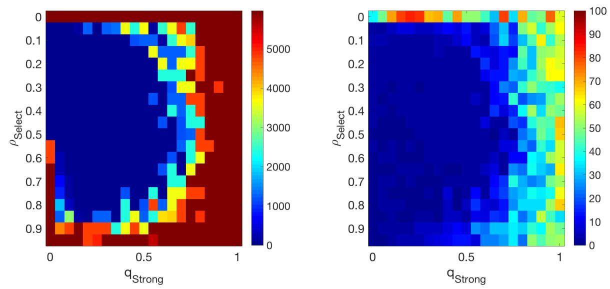
5.2 Results

We run the influence-selection dynamics with the 3-majority influence model on the initial graph defined in Section 5.1, and show the number of influence steps until consensus in Figure 5. We stop a trial if the total number of influence step is more than the two times the square of the size of the graph which is 17,844,338. The setting of bit map is similar to Figure 2, but there are 20 parameters ρ_{select} ranging from 0 to 0.95 with even space, and 21 parameters of q_{strong} ranges from 0 to 1 with even space.

Small $\mathbf{q}_{\text{strong}}$. We first consider the case where q_{strong} is small (the left part of the plots). When $\rho_{\text{select}} = 0$ the dynamics almost always time out and the number of switches is high which indicates influence may be not enough for the system to consensus when the graph has a rich structure. Interestingly, when $0 < \rho_{\text{select}} < 0.5$ (upper-left quadrant except for the top boundary), the processes reach consensus quickly, as the weak ties help the opinions to mix. This result shows moderate selection encourages agents to form (random) connections and helps the system mix. However, when selection is dominantly taken, $\rho_{\text{select}} > 0.5$ (lower-left quadrant), the processes often time out, as the selection process creates local community structures by the weak ties that hinder communication between agents of different opinions, preventing the opinions from mixing.

[2] Nodes with fewer than 10 neighbors are iteratively removed.

[3] The Jaccard similarity between u, v defined as $J(u, v) = \frac{|N_u \cap N_v|}{|N_u \cup N_v|}$, where N_v is the set of vertices adjacent to node v. The Jaccard coefficient is commonly used to measure the strength of an edges [26].



Consensus time number of switches

Figure 5: Consensus time in Facebook and number of switches before consensus. The color at each point $(\rho_{\text{select}}, q_{\text{strong}})$ **in this** 21×20 **bitmap represents the average number of influence steps before consensus (or timeout) of** 5 **trials of the dynamics** Sel-Inf$(f_{\text{inf}}, p_{\text{select}}, q_{\text{strong}})$ **with** 3**-majority measured in influence steps.**

Large $\mathbf{q}_{\text{strong}}$. In the right part of the plots with large q_{strong}, the processes often reach timeout. This may due to the community structures in strong ties of the real graph.

Interestingly, in the region of a medium-high selection rate (center height of the plots), the processes times out if the graph mostly consists of either weak ties or strong ties when q_{strong} is near 0 or 1, because of structures in strong ties and weak ties. However, if q_{strong} is near 1/2, the graph has a mixture of strong and weak ties. The community structures within the strong and weak ties seem to override each other, and so the processes reach consensus fast. This suggests multiple independent community structures help the processes reach consensus, even if individually, the community structures would stifle agreement.

The results of the simulation on the real-world graph and the synthetic one are similar when q_{strong} is small and p_{select} is large. This is not surprising because the initial condition does not matter in the above condition. When q_{strong} is large or p_{select} is small the initial graph matters a lot. Our real-world social network has 10-20 rather distinct communities, but our spatial networks, Newman Watt's model, are more uniform. Because of this, the processes on real-world network become stuck substantially more often.

6 CONCLUSION

As discovered by [11], the strength of weak ties is to get new information and fresh ideas into the comfort zone created by strong ties. However, in a time-evolving spatial network, especially one where selection happens at a substantially higher rate than influence, the role of strong ties and weak ties, in terms of spreading fresh ideas, are swapped. The weak ties are too fragile, and the power of spreading information diminishes. The selection causes the forming of weak ties that only repeat and reinforce the same opinion that the person already holds, which ironically, does not bring any new thoughts. It is nevertheless the strong ties that hold the network together, prevent it from being fully divided, and motivate the participants to compromise.

REFERENCES

[1] HA Antosiewicz. 1963. Ordinary Differential Equations (G. Birkhoff and GC Rota). *SIAM Rev.* 5, 2 (1963), 160.

[2] Eytan Bakshy, Solomon Messing, and Lada A Adamic. 2015. Exposure to ideologically diverse news and opinion on Facebook. *Science* 348, 6239 (2015), 1130–1132.

[3] Luca Becchetti, Andrea Clementi, Emanuele Natale, Francesco Pasquale, and Luca Trevisan. 2016. Stabilizing consensus with many opinions. In *Symposium on Discrete Mathematics.* 620–635.

[4] Bill Bishop. 2009. *The Big Sort: Why the Clustering of Like-minded America is Tearing Us Apart.* Houghton Mifflin Harcourt.

[5] Leticia Bode. 2016. Pruning the news feed: Unfriending and unfollowing political content on social media. *Research & Politics* 3, 3 (2016).

[6] Robin Cohen, Alan Tsang, Krishna Vaidyanathan, and Haotian Zhang. 2016. Analyzing opinion dynamics in online social networks. *Big Data & Information Analytics* 1, 4 (2016), 279–298.

[7] Willem De Koster and Dick Houtman. 2008. 'STORMFRONT IS LIKE A SECOND HOME TO ME' On virtual community formation by right-wing extremists. *Information, Communication & Society* 11, 8 (2008), 1155–1176.

[8] Richard Durrett, James P Gleeson, Alun L Lloyd, Peter J Mucha, Feng Shi, David Sivakoff, Joshua E S Socolar, and Chris Varghese. 2012. Graph fission in an evolving voter model. *PNAS* 109, 10 (6 March 2012), 3682–3687.

[9] Juliet F Gainsborough. 2005. Voters in context: Cities, suburbs, and presidential vote. *American Politics Research* 33, 3 (2005), 435–461.

[10] Mark Granovetter. 1983. The strength of weak ties: A network theory revisited. *Sociological theory* (1983), 201–233.

[11] Mark S Granovetter. 1973. The Strength of Weak Ties. *Am. J. Sociol.* 78, 6 (1973), 1360–1380.

[12] Richard A Holley and Thomas M Liggett. 1975. Ergodic theorems for weakly interacting infinite systems and the voter model. *The annals of probability* (1975), 643–663.

[13] Petter Holme and M E J Newman. 2006. Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* 74, 5 Pt 2 (2006).

[14] Nicholas A John and Shira Dvir-Gvirsman. 2015. "I Don't Like You Any More": Facebook Unfriending by Israelis During the Israel–Gaza Conflict of 2014. *J. Commun.* 65, 6 (1 Dec. 2015), 953–974.

[15] Jillian J Jordan, David G Rand, Samuel Arbesman, James H Fowler, and Nicholas A Christakis. 2013. Contagion of Cooperation in Static and Fluid Social Networks. *PLoS One* 8, 6 (2013).

[16] David Kempe, Jon Kleinberg, Sigal Oren, and Aleksandrs Slivkins. 2016. Selection and influence in cultural dynamics. *Network Science* 4, 1 (2016), 1–27.

[17] Funda Kivran-Swaine, Priya Govindan, and Mor Naaman. 2011. The impact of network structure on breaking ties in online social networks: unfollowing on twitter. In *SIGCHI.*

[18] Daniel Z Levin and Rob Cross. 2004. The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management science* 50, 11 (2004), 1477–1490.

[19] Kevin Lewis, Marco Gonzalez, and Jason Kaufman. 2012. Social selection and peer influence in an online social network. *Proc. Natl. Acad. Sci. U. S. A.* 109, 1 (3 Jan. 2012), 68–72.

[20] Ben Light and Elija Cassidy. 2014. Strategies for the suspension and prevention of connection: Rendering disconnection as socioeconomic lubricant with Facebook. *New Media & Society* 16, 7 (24 July 2014), 1169–1184.

[21] Julian J McAuley and Jure Leskovec. 2012. Learning to Discover Social Circles in Ego Networks.. In *NIPS*, Vol. 2012. 548–56.

[22] Seth C Mckee. 2007. Rural voters in presidential elections. In *Forum-A Journal of Applied Research in Contemporary Politics*, Vol. 5.

[23] Solomon Messing and Sean J Westwood. 2012. Selective Exposure in the Age of Social Media: Endorsements Trump Partisan Source Affiliation When Selecting News Online. *Communic. Res.* 41, 8 (31 Dec. 2012), 1042–1063.

[24] Elchanan Mossel, Joe Neeman, and Omer Tamuz. 2014. Majority dynamics and aggregation of information in social networks. *Autonomous Agents and Multi-Agent Systems* 28, 3 (2014), 408–429.

[25] M E J Newman and D J Watts. 1999. Renormalization group analysis of the small-world network model. *Phys. Lett. A* 263, 4 (6 Dec. 1999), 341–346.

[26] J-P Onnela, Jari Saramäki, Jorkki Hyvönen, György Szabó, David Lazer, Kimmo Kaski, János Kertész, and A-L Barabási. 2007. Structure and tie strengths in mobile communication networks. *Proceedings of the national academy of sciences* 104, 18 (2007), 7332–7336.

[27] Pew. 2014. Political polarization & media habits. (2014).

[28] Bryce Ryan and Neal C Gross. 1943. The diffusion of hybrid seed corn in two Iowa communities. *Rural sociology* 8, 1 (1943), 15.

[29] Grant Schoenebeck and Fang-Yi Yu. 2018. Consensus of Interacting Particle Systems on Erdös-Rényi Graphs. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms.* SIAM, 1945–1964.

[30] Christopher Sibona. 2014. Unfriending on Facebook: Context Collapse and Unfriending Behaviors. In *International Conference on System Sciences.* 1676–1685.

[31] Christian Steglich, Tom A B Snijders, and Michael Pearson. 2010. 8. Dynamic Networks and Behavior: Separating Selection from Influence. *Sociol. Methodol.* 40, 1 (15 June 2010), 329–393.

[32] David Strang and Sarah A Soule. 1998. Diffusion in organizations and social movements: From hybrid corn to poison pills. *Annual review of sociology* 24, 1 (1998), 265–290.

[33] Steven H Strogatz. 2014. *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering.* Westview press.

[34] Cass R Sunstein. 2009. *Republic.com 2.0.* Princeton University Press.

[35] Bo Xu, Yun Huang, Haewoon Kwak, and Noshir Contractor. 2013. Structures of Broken Ties: Exploring Unfollow Behavior on Twitter. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work (CSCW '13).* 871–876. https://doi.org/10.1145/2441776.2441875