Cracks in the echo chamber? Limits of information sorting in polarized networks

Julia Kamin[1]

APSA 2016

[1] Thanks go to countless colleagues and friends for helping me wrap my brain around the questions in this paper, in particular Skip Lupia, Stuart Soroka, Kentaro Toyama, Zander Furnas and Joe Ornstein, with special thanks to the ever patient Scott Page and Nicholas Valentino.

Cracks in the echo chamber? Social media & the limits of ideological information sorting

In today's partisan political climate, social media is often implicated for increasing polarization by creating echo chambers that insulate ideologues from challenging political information. There is evidence, however, that online social networks may not amplify the sorting of political information; rather, they may expose users to more ideologically diverse information than they would see via other media (Flaxman et al, 2016). Recent research on Twitter and Facebook suggests this could be the case by virtue of the relative diversity of our "egonets," those we directly connect to on social network sites. (Barbera et al, 2015; Bakshy et al, 2012). In this project, I examine information diffusion in networks beyond egonets and find another reason we might expect social networks to mitigate rather than exacerbate political information sorting. A common intuition is that as ideologically charged information flows through a network it will tend to concentrate among users who share that ideology. Using mathematical and agent-based models, I demonstrate somewhat counter-intuitively that in general we should expect the diffusion of ideological information to result in a news distribution that is more ideologically balanced than the connections of the people in the network themselves. At the same time, while diffusion in networks may not lead to high levels of information sorting, it may still favor the spread of information for which ideologues are strongly biased. That is, diffusion in networks results in a relative balance of information, but it may be a balance of ideological talking points.

Keywords: social media; polarization; agent-based models

Introduction

Social media is changing the way we receive news, with consequences not just for the quality of news we consume - but also for how polarized our news environments become. Political scientists have been wary of citizens' tendency to consume information that aligns with their beliefs since the middle of the last century (Lazarsfeld et al, 1968; Sears & Freedman, 1967). The advent of cable and the internet, which gave citizens latitude to select news tailored to their views, only raised concerns that the divide in our political information diets would increase, eventually posing a threat to political discourse and even the legitimacy of the democratic process itself (Stroud, 2007; Sunstein, 2009). Today, a common impression is that social media continues to amplify that divide: as we increasingly receive news from our like-minded friends online, we narrow our scope as individuals and collectively draw ourselves deeper into ideological echo chambers (Sunstein, 2009; Manjoo, 2011).

Research that quantifies the level of ideological polarization on online social networks seems to support this view. Whether you slice networks by how we friend, follow, hashtag, reply or like, graphs of social networks consistently show liberals clustering with liberals and conservatives with conservatives, with the like-minded sharing like-minded information (Boutyline & Willer, 2013; Conover et al, 2012; Hanna et al, 2011; Himelboim et al, 2013).

Yet, while there is little doubt that social media sites expose us to information that aligns with our views, it is still unclear if social media makes information sorting *worse* than it would be in its absence. That is, with citizens newly able to receive political information from online social networks, are they exposed to more homogenous information that aligns with their beliefs or, rather, to a greater diversity of information than they would be exposed to if social media did not exist? Likewise, are the sets of information that liberals and conservatives are exposed to with social media more distinct than what they would otherwise consume?

There are a few reasons social media watchers conjecture online social networks increase information sorting: for one, we tend to connect with like-minded friends; those friends may be ideologically biased in the information they share; and, finally, because our friends likewise connect to similar others, collectively we will form tightly knit echo chambers that amplify our biases in who we link to and what we post.

Several scholars have begun to question the assumptions above, however, focusing in particular on whether our online networks are as homophilous as feared (Flaxman et al, 2016; Barbera et al, 2015; Bakshy et al, 2012). In this paper I extend that work by testing the assumption that the diffusion of political information in networks contributes to information sorting. If we think of an information network in which liberals tend to cluster with liberals and conservatives with conservatives, and we assume those individuals are heavily biased in what they share, a common intuition is that those two biases would amplify each other, further concentrating messages into ideological cul-de-sacs as they diffuse through the network. Yet, the degree to which message diffusion in polarized networks results in greater information sorting has not been examined.

In this paper, I take a preliminary step in measuring possible diffusion effects on sorting by using both mathematical and agent-based models that simulate message diffusion across a variety of networks. I ask, given different assumptions about our connections and our micro decisions about what to post, *should we expect the process of information diffusion on networks to expose users to more or less information diversity than the diversity of their friends in the network*? Likewise, but at the macro-level, *will the diffusion of political information through networks result in more information sorting, where liberals and conservatives are exposed to more distinct sets of information?* In short, do our *micro*-level biases on who we connect to and what we share actually translate into *macro*-level sorting that is even more imbalanced, as many fear?

Across a range of network types and assumptions about diffusion patterns I find the surprising result that diffusion produces information sorting that is *less* divided (more evenly distributed) than the connections of people who make up the network. Likewise, users are exposed to more diverse (less homogenous) information than the diversity of their connections. For example, in a network where 80% of users' friends share their ideology, political information will always be less than 80% concentrated within one ideological type and users will likewise be exposed to less than 80% congruent information (or, conversely, exposed to more than 20% cross-ideological information). However, models also suggest that our intuition that social media polarizes political information is not completely awry; although diffusion results in a more balanced distribution of ideological messages, models suggest that networks may have the effect of sorting *out* information for which there is not a strong ideological signal. That is, while networks may lead to more balanced information, it may be a balance of strongly ideological talking-points.

Information sorting pre and post social media

Before discussing the risks of political information sorting, it should be defined: understood at the network – or population – level, "political information sorting" is the state in which liberal information tends to be more concentrated among liberal individuals (and conservative information with conservatives); it can also be the process by which information sorting increases. At the individual level, information sorting would result in individuals being exposed to a greater proportion of like-minded information and less diverse information.

Political theorists give us a number of reasons a democratic society should be concerned about political information sorting. For one, as J.S. Mill argues, a democracy benefits when its citizens are aware of beliefs that challenge their own; it is only in the "collision of adverse opinions" that society as a whole can move in the direction of truth - and so be better able to understand and address the problems it faces (Mill, 1869). Awareness of opposing views also legitimates the democratic process by exposing voters to the views - and reasons for those views

- of opposition leaders; in doing so the odds are increased that supporters of losing parties view the outcome of elections as legitimate (Manin et al, 1987; Sunstein, 2009).

This ability to tolerate the opposing party may seem like a meager aspiration, but it is essential for democracy. In the US today, indeed, we may have cause to be particularly concerned about tolerance; although Americans may not be polarizing on issues (Fiorina & Abrams, 2008), affective polarization - negative feelings towards the opposing party - is on the rise (Iyengar et al, 2012). Exposure to cross-ideological information has the potential of not only of diminishing affective sorting (Parsons, 2010; Garrett et al, 2014) but of promoting higher levels of political tolerance (Mutz, 2002; Mutz & Mondak, 2006; Erisen & Erisen, 2012; Ben-Nun Bloom & Levitan, 2011).

In spite of the virtues of exposure to challenging political information, it has long been observed that Americans by and large tend to do the opposite, consuming news that aligns snugly with their political beliefs (Lazarsfeld et al, 1968; Sears & Freedman, 1967; Hart, 2009). Why that is so has been the subject of much research. Before cable and the Internet, to the extent political information sorting existed, it was likely a result of “de facto selective exposure” (individuals’ news consumption habits and their “incidental” exposure to news in their environment) and the “two step flow” of information, in which “opinion leaders” relay information to those in their community. (Sears & Freedman, 1967; Lazarsfeld & Katz, 1955). Homophily – our tendency to associate with others who share our views – insures that both incidental exposure and the “two step flow” deliver information that aligns with our political leanings. The arrival of cable and the Internet gave rise to a potential additional sorting force – “active selective exposure”; consumers could now more easily choose news to suit their political perspective – an ability that raised concerns information sorting would rise to destabilizing levels

(Negroponte, 1996; Sunstein, 2009). Evidence that those fears have borne out is, however, mixed (Hart et al, 2009; Gentzkow and Shapiro, 2011; Stroud, 2007; Prior, 2005).

To understand information sorting on social media we must likewise sort out the mechanisms at work. Homophily continues to play a – perhaps even greater - role in determining the information we see in online social networks; we not only connect to people in our lives who tend to share our views (Huckfeldt & Sprague, 1987), but we may also seek out and connect to strangers who share our interests and worldview (Bimber, 1998). What plays a diminished part, however, is active selective exposure; the set of news stories we are exposed is less a function of our choices, but rather the choices of our friends. Our friends' "selective sharing," then, replaces "active selective exposure." Finally, the stories we see is not just determined by our direct connections and their choices on what to post, but also what our connections are exposed to via their friends. In sum, in social media information sorting is a result of homophily (who we connect to), selective sharing (the decisions those connections make on what to share) and the broader network (who our connections are connected to and the resulting macro- patterns of diffusion).

A common view is that these forces will combine to strengthen echo chambers. As the Encyclopedia of Social Media and Politics describes it, "Using social network sites… individuals strengthen the echo chambers by reposting media content and offering supporting materials…. consequently, users end up hearing the same information over and over" (see also Sunstein, Pariser and Manjoo). And much evidence supports the view that we do tend to connect with people who share our ideology and share like-minded information on social media. Studies that chart networks on Twitter, for example, find that we consistently cluster by political views, regardless of whether you are looking at those connections by who retweets whom (Conover, 2012), hashtag use (Hanna et al, 2011), key words (Himelboim et al, 2013) or who follows political accounts (Boutyline & Willer, 2015). (See also: Feller et al, 2011; Gruzd & Roy, 2014; Lerman & Ghosh, 2010; Yardi & Boyd, 2010.)

Yet researchers are beginning to find evidence that the sorting effect of social media is not as bad as we imagine – and that social media even may lead to less sorting than other media. Flaxman and colleagues (2016), for example, compare the diversity of news stories users consume via social media to the stories they read via other online entry points; they find that although the average liberal and conservative are more ideologically divided in the news they consume via social media (compared to the news they access by directly going to news sites), at an individual level users read a more diverse set of news when directed to stories by social media. Barbera and colleagues (2015) offer one explanation for why that might be the case; they investigate levels of homophily on Twitter and find that most of us tend to have more diverse friend connections on Twitter compared to the new organizations we follow. Researchers at Facebook similarly find that we connect to a surprising number of “weak ties” (connections outside of our tightly knit group of friends) and that those ties have a disproportionate influence on the stories we repost to share with our friends (Bakshy et al, 2012).

But while there is evidence that homophily and “selective sharing” may not be as rampant as often imagined, we know little about the process of diffusion on social networks and whether it will add to or mitigate information sorting. In the modelling world, researchers have looked extensively at polarization in networks but only in terms of how beliefs are influenced (Baldassarri & Bearman, 2007; Banisch et al, 2012; Bednar et al, 2010; Dandekar et al, 2013; Flache & Macy, 2011). Others have also looked at the diffusion of information, but not in the context of ideological sorting (Centola & Macy, 2007; Goldenberg et al, 2001). Finally, Siegel

(2013) has used models to demonstrate how social networks can amplify media bias, but in that work bias represents a non-status-quo position and information moves through a non-polarized network. No work to date has examined how diffusion of information in a polarized network will affect the ultimate level of political information sorting in social networks.

Network effects: how network diffusion influences information sorting

To examine the impact of diffusion on information sorting in a polarized network, I model diffusion on a series of networks, using both mathematical calculations and agent-based models to estimate expected levels of sorting. Models help us understand the dynamics of diffusion and sorting not by telling us how the world actually works, but by showing how, *given our assumptions*, we would *expect* the world to work. They are especially useful in understanding complex, nonlinear, interactions where our intuition may fail us.

In this case, our intuition might tell us that, if social media users are biased in who they connect to and what they share, then those biases would have an amplifying effect on the sorting of political information. That is, if conservatives in a network tend to connect with fellow ideologues and also have a bias toward reposting like-minded information, we might guess that as ideological messages move through a network they would become increasingly concentrated among sympathetic ideologues. Similarly, we might intuit that in such a network, conservatives and liberals individually would receive a greater concentration of ideologically-friendly information than the level to which they "selectively link" to like-minded friends.

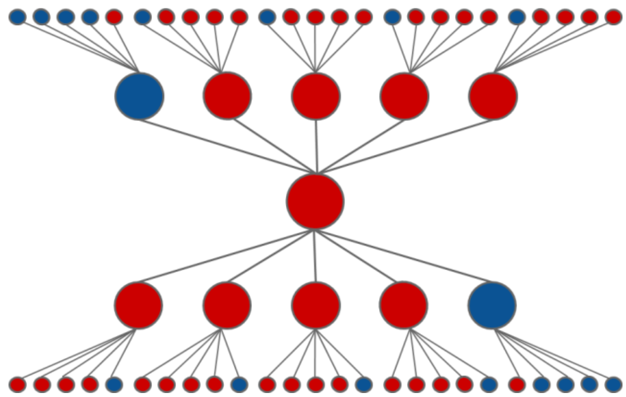
The models in this paper test those intuitions by simulating the diffusion of ideological posts in various polarized networks. In each model, individuals have a given ideology, and are inclined to "friend" like-minded others with a given bias, *F*, while eschewing counter-ideologues with an inverse biase (*1-F*). Users likewise post and re-post political messages with a bias toward messages that are similar to their ideology (*R*) and against messages that run counter to their ideology (*1-R*). As the model simulates diffusion of messages through the network, I capture levels of sorting with each wave – or time period – to see whether it matches our intuition.

As with any model, these set aside many factors that may be relevant to understanding the distribution of information in *real* social networks. For one, they are static; I do not model the "friending" or "defriending" process on social networks which will have consequences on what information we are exposed to. The models in this paper also do not include third party algorithms that influence which posts individuals are exposed to; individuals "see" all information that their friends post. Finally, models do not consider the degree to which individuals give attention to or are influenced by the posts they are exposed to. Ultimately we should be interested in how political messages affect beliefs and behavior, but these models only address the first step in influence: exposure.

Infinite models

To create the rawest understanding of the degree to which networks amplify political information sorting I first model a directed infinite network, one in which all individuals ("nodes") have links ("edges") flowing in (from which they receive messages) and links going out (to which they forward messages), with no cross-over in links between nodes. In such a network, there are an infinite number of nodes like the central one shown in Figure 1, each with an infinite tree where messages flow from and likewise an infinite tree where that nodes' messages flow out.

Figure 1. A slice of an infinite tree model, with two degrees shown.



In this simple abstract model, each node is either liberal or conservative and all nodes are assigned the same level of "selective linking" to friends (*F*) and "selective reposting" (*R*). "Selective linking" is one's bias in connecting to like-minded nodes. In a network with a selective linking rate of 0.8, like the one in Figure 1, 80% of one's links will be with nodes that share one's ideology. "Selective reposting" is a node's bias in deciding what information to re-share: a selective reposting bias of 0.9 would mean that, if a liberal node received an equal number of liberal and conservative messages, 90% of those it would forward would be liberal messages. Finally, to keep the model – and our calculations – simple, each node has the same number of friends it "follows" and is followed by.

Looking down the tree: How concentrated do ideological messages get?

Using this model we can first answer the question: as conservative (liberal) messages move through the network, do they become more or less concentrated among conservatives (liberals)? In other words if a conservative node posts a conservative message, as each successive wave of nodes choose to repost that message, is that message seen by proportionally more or fewer conservative nodes?

To answer that question, let us define a “wave” as one round of reposting (or “infection” to use the language of contagion) and “exposure” to a message. If we then take a conservative message and send it out from an initial conservative node, at each wave the number of newly exposed conservatives and liberals (Ce and Le) in a given wave (or time period, *t*) will be a function of the number of newly infected conservatives and liberals (*Ci* and *Li*) in that wave, how many friends they have, and the level of selective linking. In a network where all nodes have 5 friends and have a selective linking bias of 0.8:

$$Ce_t = 0.8 * 5 * \ Ci_t + 0.2 * 5 * \ Li_t \qquad (1a)$$

$$Le_t = 0.2 * 5 * \ Ci_t + 0.8 * 5 * \ Li_t \qquad (1b)$$

In the case above where we start off with one conservative posting a conservative message, so where Ci_t = *1* and Li_t = *0*, the number of exposed conservatives, Ce_t, will be 4 and the number of exposed liberals, Le_t, will be 1.

Since ultimately we are interested in the ratio of conservatives to liberals, we can drop the number of friends each node has (in this case “5”) and more broadly say that for each wave:

$$Ce_t = \ Ci_t * F + \ Li_t * (1 - F) \qquad (2a)$$

$$Le_t = \ Ci_t * (1 - F) + \ Li_t * F \qquad (2b)$$

In this example, *F* is set from the perspective of a conservative, so the number of conservatives who are exposed to a conservative message is both a function of the number of conservatives who post a message (Ci_t) times conservatives’ bias toward linking to other conservatives (*F*) plus the number of liberals posting a conservative message times the likelihood of a conservative linking to a liberal (*I-F*). Liberals will be exposed in opposite proportions: the likelihood they will connect to a conservative is *1-F*, and to a liberal *F*.

Once those conservatives are exposed to the conservative message, the number who decide to repost a message for the next wave of diffusion is a function of how many conservatives are exposed times their likelihood of reposting that message (R). (And liberals at a rate of (*1*-*R*)):

$$Ci_{t+1} = R * [\,Ce_t] \qquad (3a)$$

$$Li_{t+1} = (1 - R) * [\,Le_t] \qquad (3b)$$

Putting the two sets of equations above together, we can represent the number of conservatives and liberals newly exposed to a conservative message as a function of those exposed in the previous wave.

$$Ce_{t+1} = F * R * [\,Ce_t] \ + \ (1 - F) * (1 - R) * [\,Le_t] \qquad (4a)$$

$$Le_{t+1} = (1 - F) * R * [\,Ce_t] \ + \ F * (1 - R) * [\,Le_t] \qquad (4b)$$

Where, again:

- *Ce* and *Le* = the number of Conservatives and Liberals “exposed” to a message.
- *Ci* and *Li* = the number of Conservatives and Liberals “infected” by a message, i.e. who repost the message.
- $F = Selective\ Linking$, or users’ bias toward linking to co-ideologues.
- $R = Selective\ Reposting$, or users’ bias in reposting messages that align with their ideology.

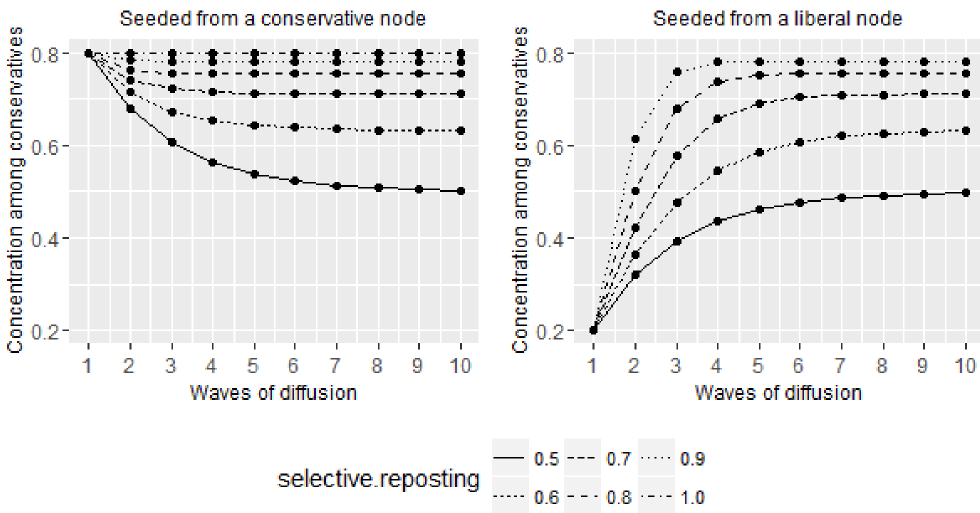
To find the proportion of newly exposed nodes that are conservative, we calculate the number of conservatives who were exposed the message (Ce_{t+1}) and divide that by the total number of nodes exposed ($Ce_{t+1} + Le_{t+1}$):

$$\frac{Ce_{t+1}}{Ce_{t+1}+Le_{t+1}} \qquad (5)$$

Staying with the example network above, in which nodes have rates of 80% selective exposure and 90% selective sharing, we can take the proportion of nodes exposed to the initial message that are conservative (4/5) and liberal (1/5) and plug those proportions in to calculate levels of sorting for the second wave. In doing so we see that 78.4% of those exposed will be conservative. To find the proportion of nodes exposed that are conservative in the next wave, we repeat the process and see the percentage of exposed that are conservative drop even more slightly to 78.2%. By the fourth wave, that number remains; we appear to have reached an equilibrium of 78.2%.

If we iterate through the same process for other levels of *R*, selective reposting, in an infinite network with a selective linking rate of 0.8, each time we likewise arrive at an equilibrium, as seen in Figure 2a. The first thing we may notice is that for all levels of selective sharing except for 1.0 we reach an equilibrium which is somewhere between 0.5 and the network's level of selective linking. Only when nodes in the network use a 100% / 0% formula for sharing – i.e. sharing only messages that align with one's ideology – does the sorting of information exposure remain at 80%. *For all other levels of selective sharing, the distribution of the message reaches an equilibrium that is less sorted than the people in the network*. At the other end of the reposting spectrum, where nodes are indiscriminate in which messages they pass along – i.e. they pass all along at a 50% rate – sorting disappears.

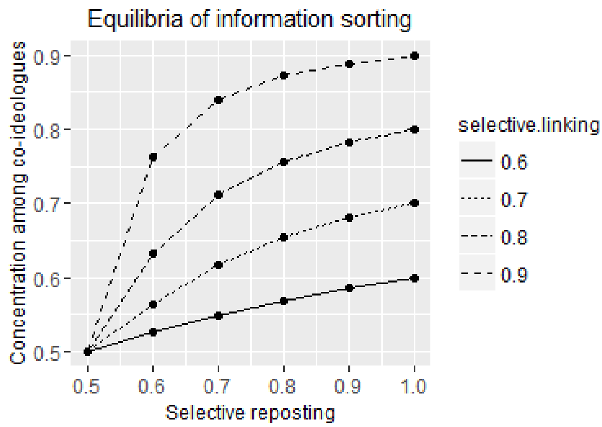
Figure 2a and 2b. Proportion of nodes newly exposed to a conservative message that are conservative at each wave of diffusion, at varying levels of Selective Reposting.



The above equilibria were reached when a conservative node was the originator of a conservative message. But a true equilibrium should be reached regardless of the starting point. As a check, we run our calculations over, this time with a liberal node as the initiator of a conservative message and find, indeed, the same equilibria are arrived at, as shown in Figure 2b. While Figure 2 affirms that equilibria are reached regardless of the starting point, it also demonstrates that when we speak of the *process* of information sorting, the *starting point* matters. If we start with a conservative node sharing a conservative message, sorting *decreases* from 80% to 78.2%. But if we start with a liberal sharing a conservative message sorting *increases* from 20% to 78.2%. As will be discussed later, the degree to which social media increases or decreases sorting will depend upon what we assume about our starting points.

To better capture the relationship between our reposting biases and how much sorting we can expect in networks at equilibria, we can graph the equilibria reached at varying levels of R and F, as seen in Figure 3. Regardless of the selective linking rate of a network (denoted by a single line in Figure 3), at levels of 50% selective reposting, information sorting will eventually disappear – that is, over time half of those exposed to a message will be conservative and half liberal. At the other extreme, with 100% selective sharing, at equilibrium sorting will be at the network's level of selective linking. As before, we see that for all other levels of selective sharing, the *equilibria reached will be less sorted than the network's level of selective linking*.

Figure 3. Equilibria of information sorting reached for each level of Selective Reposting (R) in networks with different levels of Selective Linking (F)



Getting at the equilibria mathematically

We arrived at the above equilibria by iterating calculations for each wave. We can also, however, derive equilibria mathematically. To do so we start with the formulae derived above, which tell us the number of exposed nodes that are conservative or liberal in a given wave based on the number exposed in the previous wave:

$$Ce_{t+1} = F * R * [\,Ce_t] \;+\; (1-F) * (1-R) * [\,Le_t] \qquad \text{(6a)}$$

$$Le_{t+1} = (1-F) * R * [\,Ce_t] \;+\; F * (1-R) * [\,Le_t] \qquad \text{(6b)}$$

At equilibrium, by definition, the proportion of conservative to liberal nodes that are exposed to a message will stay the same from wave to wave. To make it simpler to find that proportion – and to generalize the equation to not refer to conservatives alone – we substitute *P* (proportion of nodes receiving a politically aligned message) for *Ce* and *1-P* (proportion of nodes receiving the same message, but for whom the message is counter-ideological) for *Le*, giving us the equations:

$$P_{t+1} = F * R * [\, P_t] \; + \; (1-F) * (1-R) * [\, 1 - P_t] \qquad (7a)$$

$$1 - P_{t+1} = (1-F) * R * [\, P_t] \; + \; F * (1-R) * [\, 1 - P_t] \qquad (7b)$$

Mathematically we know that:

$$\frac{P_{t+1}}{1-P_{t+1}} = \frac{F*R*[\, P_t] \; + (1-F)*(1-R)*[\, 1-P_t]}{(1-F)*R*[\, P_t] + \; F*(1-R)*[\, 1-P_t]} \qquad (8)$$

Since again at equilibrium $P_{t+1} = P_t$, we can denote *P* at both time periods as *P*, giving us:

$$\frac{P}{1-P} = \frac{F*R*P \; + (1-F)*(1-R)*[\, 1-P]}{(1-F)*R*P + \; F*(1-R)*[\, 1-P]} \qquad (9)$$

Solving for *P* we find:

$$0 = (2R-1) * P^2 + (2 - 2R - F) * P - (1-F)(1-R) \qquad (10)$$

We can now plug in any *F* and *R* to the equation to get the proportion of nodes that receive a politically aligned message at equilibrium; we find, indeed, the same results as seen in Figure 3.

Looking up the tree: How diverse are the messages we are exposed to?

Above, we looked at *the proportion of nodes* that would receive an ideologically-aligned message in a network given a certain selective linking and selective reposting rate. But what about the *proportion of congruent messages* that an individual would receive?

To answer that question, we start by observing that in a network one's exposure to ideologically aligned information will be a function of both of our bias toward linking to those who share our beliefs (*F*) and our friends' bias in selectively sharing information that aligns with our beliefs (R):

$$Exposure\ to\ ideologically\ aligned\ information = F * R + (1 - F) * (1 - R) \tag{11}$$

The first term $F * R$ represents the congruent information we receive from our co-ideologues, and $(1 - F) * (1 - R)$ the congruent information we receive from our counter-ideologues. But we do not stop there. Right now the R term tells us how much congruent information our friends post without regard to how much ideologically aligned information they themselves receive; it assumes they receive equal parts liberal and conservative information. But since they have the same rate of selective exposure as we do, that would not be the case; we should assume they, like us, receive more congruent information from their friends.

To calculate the relative amount of congruent information we receive taking into consideration the relative amount of congruent information our friends receive, we need to calculate the proportion of congruent information we are exposed to given how much congruent information our friends receive, which I will call *P*:

$$Exposure\ to\ ideologically\ aligned\ \ information =$$

$$\frac{F*R*P+(1-R)(1-R)(1-P)}{[F*R*P+(1-F)(1-R)(1-P)]+[F(1-R)(1-P)+(1-F)*R*P]} \tag{12}$$

The top term represents the amount of congruent information we would receive; it is the same as the term previously discussed, except adding in *P*, how much congruent information our co-ideologues receive and (*1*-*P*), how much congruent (to us) information our counter-ideologue friends receive. The bottom term represents *all* the information we would receive including congruent (same as the top term) and incongruent.

As we did earlier, we can calculate the equilibrium using the expression above and noting that, at equilibrium our exposure to congruent information will be the same as our co-ideological friends, or:

$$P = \frac{F*R*P+(1-F)(1-R)(1-P)}{[F*R*P+(1-F)(1-R)(1-P)]+[F(1-R)(1-P)+(1-F)*R*P]} \qquad (13)$$

When we solve for P, we get:

$$0 = (2R-1)*P^2 + (2-2R-F)*P - (1-F)(1-R) \qquad (14)$$

This is the same solution as previously derived when we looked at the relative amounts of conservatives and liberals who would receive a conservative message. In other words, at equilibrium the level of information sorting in a network will be the same as the degree of congruence for each node.

Taken from the perspective of the information that a node receives, we can intuitively see why our level of congruence will never be above the level of selective linking. Considering again a conservative node who has 80% conservative friends and 20% liberal, if we were to imagine that nodes were 100% biased in the information share – that is, that conservatives *only* forwarded conservative news and liberals liberal news – then our conservative node above could not possibly receive more than 80% conservative information. The only way he could is if his liberal friend was less than 100% biased in the news he shared; but if that were the case then the liberal friend’s friends would receive substantially more diverse information – and we would once again end up with slightly less than 80% information sorting.

Modeling on random and small world networks

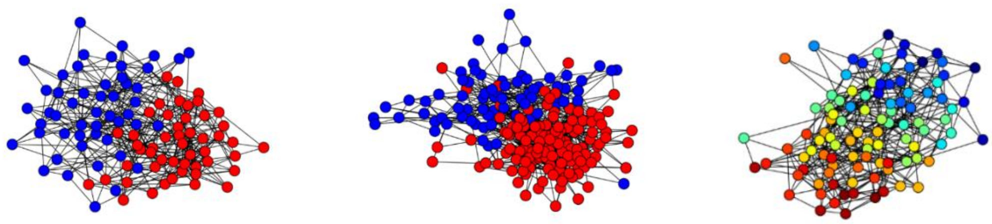
While modeling diffusion on an infinite network helps us correct our intuition on information sorting, it is as far removed from true online social networks as imaginable. To move closer to what diffusion might look like in “reality,” I construct more complex networks and use agent based models to simulate diffusion given different levels of selective linking and selective posting.

To do so I use python coding and the networkx module to create three types of networks: binary random networks, binary “small world” networks and continuous random networks.

Random networks are networks in which nodes form edges with other node with a given probability. In the binary random networks I construct, nodes are assigned an ideology (liberal or conservative) and have a bias toward connecting with nodes that share their ideology. (Figure 4a) Small world networks are distinct from random networks in that they both have high levels of clustering (your friends tend to also be friends) and as well as short paths across the network (there are few degrees of separation between any two nodes), two qualities that are found in many real-life networks, including social networks (Watts & Strogratz, 1998). (Figure 4b) Finally, because humans are rarely binary, I construct networks where nodes are not limited to a choice of “liberal” or “conservative” but take on an ideology on a continuum from 0 (extremely conservative) to 1 (extremely liberal). In these networks nodes have a bias to link with others that are ideologically close. (Figure 4c)

Figure 4. Example binary random network, binary small-world network, and continuous random network, each with a selective linking rate of 0.8.

a. Binary random network b. Binary small-world network c. Continuous random network



Simulating diffusion

Using the three types of graphs described above, I simulate the diffusion of political information in networks with different levels of selective linking and selective reposting. In each simulation:

- A seed node is randomly selected to post an initial message with a given ideology.
- That node's connections then see the message (are 'exposed') and use an algorithm to decide whether to repost that message (become 'infected'). That algorithm is based on the node's ideology and the ideology of the posted message. Exposure and infection make up one wave, or one time period.
- In the next wave, connections of the newly "infected" nodes become newly exposed and will likewise decide whether to repost the message.
- Each message moves through the network for 5 waves, unless it dies out sooner.

For every network (i.e. type of graph with a given average rate of selective linking and selective reposting), I generate 100 graphs of 1,000 nodes each, in which each node has an average of 8 links, and I simulate the diffusion of 100 messages. I make two calculations to track sorting at each wave. First, the model calculates the degree of *sorting at the network level*: for each message at each wave I calculate the percentage of all newly exposed nodes that share the message's ideology. Next, the model calculates the average *level of congruence* for all nodes at each wave: after all 100 seeds diffuse through the network, the model looks at the total number of messages that each node was exposed to at each wave and calculates the percentage of those messages that align with the node's ideology.

Results

Across all network types and decision algorithms, I find consistency with the results of our infinite model. To walk through the results, I use the example of graphs with 0.8 selective linking, in part because they demonstrate the effects most clearly, but also because 0.8 selective linking best approximates the level of bias in online social networks (Bakshy et al, 2012; Barbera et al, 2015).

Binary networks

Starting with binary networks, recall that our infinite networks showed that after a few waves of reposting, messages would reach an equilibrium level of sorting.

Comparing results from random and small world binary networks, as seen in Figure 5a, simulations show a pattern similar to the infinite network. At the lowest levels of selective reposting where R=0.5, messages reach an almost equal number of conservatives and liberals by the fifth wave of diffusion. At the other end of the spectrum, with reposting biases at 0.95, messages continue to reach high proportions of nodes that agree with the message's ideology, yet the distribution of the message never exceeds an imbalance greater than 0.8. In Figure 5b, we see similar results when looking at the congruence of messages nodes are exposed to; with each wave of reposting, the set of messages nodes receive become less strictly congruent with the nodes' beliefs, and so more diverse.

Figures 5a and 5b. Simulations run on agent-based models with selective linking levels of 0.8, at selective reposting rates from 0.5 to 0.95. On binary random and small world networks.

5a. Levels of information sorting.

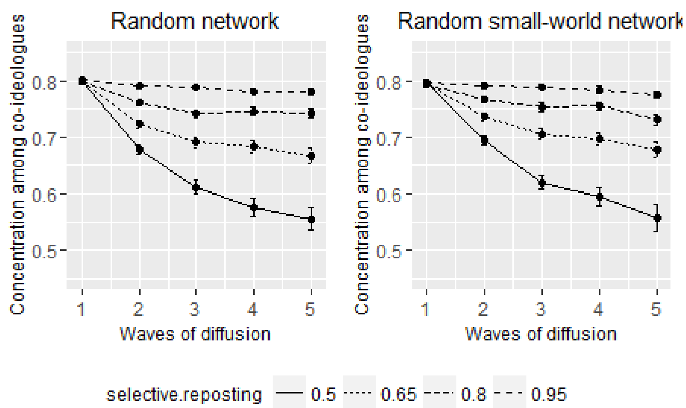
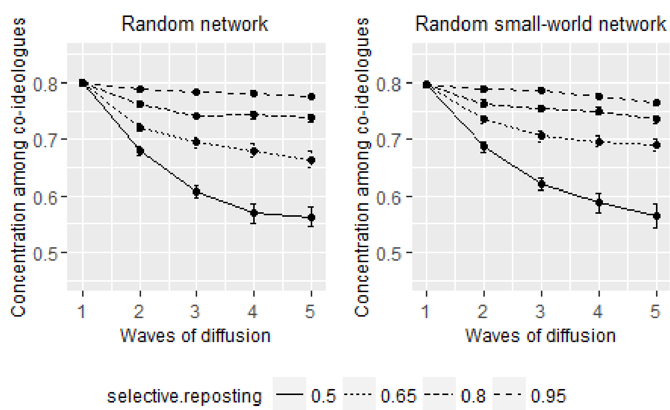


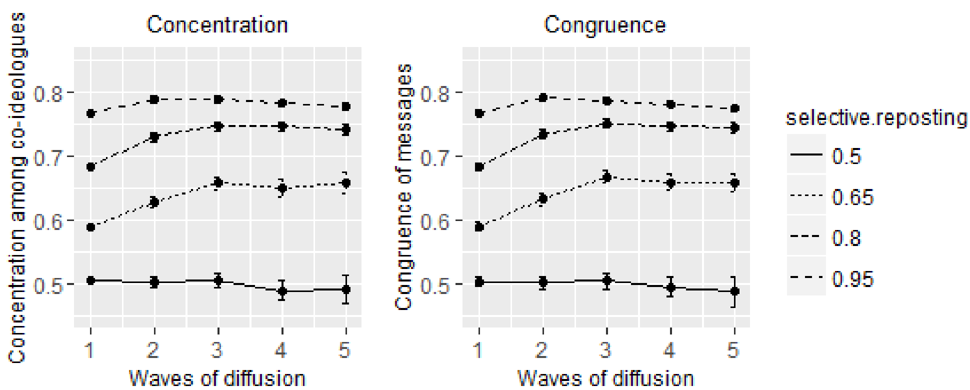
Figure 5b. Levels of congruence.



Selecting a different starting point

In the simulations above, nodes initiated posts that were always aligned with their identity. But what if nodes were not 100% biased in their initial posts, but posted with the same bias as they reposted messages? For example, a conservative node with a 0.95 reposting bias would initiate conservative messages 95% of the time and liberal messages 5% of the time. If we simulate those diffusions we see, in Figure 6, that sorting and congruence *increases* from the first wave. As discussed earlier, when discussing the process of information sorting, starting points matter. Yet if we are interested in outcomes, we notice that – regardless of starting point – the distribution of ideological posts by wave 5 tends toward the same level, a level which is – again – less sorted than the nodes themselves.

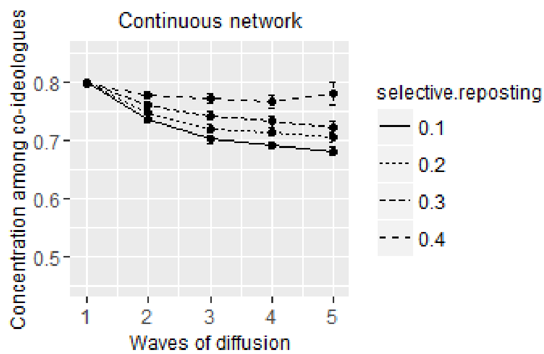
Figure 6. Information sorting and congruence in the diffusion of messages through Binary Random networks, with initial posts reflecting nodes' reposting biases.



Continuous networks

A similar pattern is seen in continuous networks, even when we make strong assumptions about bias in reposting. I ran simulations in which nodes are both biased to repost messages with ideologies that are closer them (measured by the likelihood that a message's ideology falls in a normal distribution around the node's ideology) and a node's bias is relative to the extremity of their ideology. Looking at varying levels of how "close" a message needs to be for a node to be biased towards it, using standard deviations of 0.1, 0.2, 0.3 and 0.4, in Figure 7 we see political information in general becomes more balanced over time. In the case where nodes are biased toward messages with a standard deviation of 0.1 ideological points away, however, we see signs of that trend reversing; this exception may indicate that there are outlying assumptions in which greater information sorting may occur.

Figure 7. Information sorting in random continuous networks, with average selective linking levels of 0.8. In these simulations nodes' reposting biases are relative to the extremity of their ideology, with biases applied to messages closer to the node's ideology (normally distributed with standard deviations of 0.1 to 0.4).



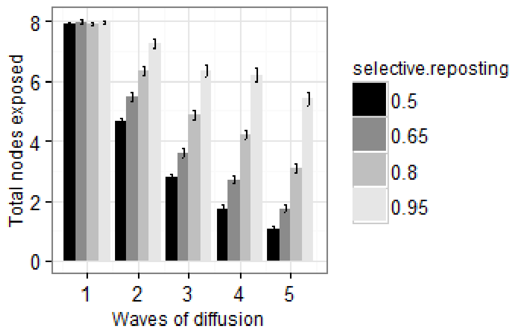
Another mechanism: sorting out neutral information?

Although models tell us that in general networks have a moderating effect on information sorting (or at least keep sorting at levels lower than rates of selective linking), they may – paradoxically - also have the effect of favoring ideologically biased information over neutral information.

In the simulations run above, our results show the *proportion* of liberal to conservative nodes that are exposed to a message in any given wave (and likewise the proportion of ideologically congruent messages individuals are exposed to). If we examine the *absolute* number of all nodes that are exposed (or messages received), however, we see that when nodes on average have an extreme ideological bias in their reposting, messages proliferate to a far greater degree (among both liberals and conservatives).

For example, if we consider a network with selective linking bias of 0.8 and look at the number of nodes exposed to messages when there is a selective reposting bias of 0.5, 0.65, 0.8 and 0.95, we see in Figure 8 that when the bias is 0.95, by the fifth wave political messages are being shared at five times the rate of messages when there is a 0.5 bias.

Figure 8. Number of nodes exposed to a message at each wave of diffusion, in binary random networks with a selective linking rate of 0.8, at rates of selective reposting from 0.5 to 0.95.



Before deriving any insights from that result, we should note that these biases are biases toward the ideology of the message only – and are devoid of other biases, for example a bias towards message quality or sensational value. But if we assume no other preferences other than a bias toward ideologically-favorable information, then there are two possible insights we may have. First, we might believe that in networks where individuals are highly ideologically biased in the information they repost, that those networks will have proportionately more ideologically biased posts *compared to other networks* where there is little bias. We may also compare how much an ideologically divisive message would spread *compared to another message* that is neutral message *in the same network.* That is, if we imagine the same network has a 95% bias when it sees an divisive post from Rachel Maddow or Glenn Beck, but a 50% bias when it comes to an politically neutral article from the USA Today, we would expect Rachel and Glenn to proliferate and dwarf the number of USAT articles that diffuse through the network.

In this way, diffusion through networks may not lead to extreme levels of sorting, but it may indeed polarize the information that circulates in networks by favoring the proliferation of extreme messages over moderate ones.

Discussion

In this paper, I used mathematical and agent-based models to test the intuition that social networks – by layering our tendency to link to like-minded friends on our tendency to repost ideologically congruent messages – will compound those biases to create even deeper ideological echo chambers. Simulations on those models show the counter-intuitive finding that the diffusion of ideological messages through a polarized network will tend toward an equilibrium that is more balanced than the connections of the people in that network. So in a network where, on average, 80% of individuals' friends share their ideology, no matter how biased users are in the information they share, users will always be exposed to more than 20% incongruent information and, likewise, the audience of politicized messages will be more than 20% ideologically opposed to that message.

These findings are consistent across a number of network types, including random binary networks, small world binary networks and random continuous networks. At least one research project suggests these findings may not be limited to hypothetical models. A 2015 Facebook study finds that although only 20% of liberals' friends are conservative, and 18% of conservatives' friends liberal, liberals and conservatives alike are exposed to more counter-ideological information than we might expect: 24% of the information liberals receive is conservative, while for conservatives, the amount of cross-cutting information is a remarkable 38% (Bakshy et al, 2015).

We might also expect this finding to be near universal, by realizing that even in the most biased network – where individuals exclusively share congruent messages – users will be exposed to a proportion of congruent messages that matches the proportion of friends that share their view, but not more. Yet it may be there are outlier network structures and parameters that would result in information sorting levels more extreme than the networks themselves. Further research is warranted to find those conditions. It would also be valuable to explore network attributes that may lead to relatively smaller degrees of sorting and so more diversity, such as network density, clustering coefficients, etc.

Finally, while models tell us that the diffusion of political information in a polarized network will not result in run-away information sorting, they do suggest that network diffusion may favor messages that have a strong ideological bias. That is, although news posts from HuffPo and Fox may be relatively evenly spread across a polarized network, the dynamics of diffusion will favor the diffusion of those posts over less ideologically divisive news stories from USA Today or Time. Further research is called for in understanding what types of political news stories – from both left and right – we should expect to proliferate even more across a network.

References

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130-1132.

Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012, April). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web* (pp. 519-528). ACM.

Baldassarri, D., & Bearman, P. (2007). Dynamics of political polarization. *American sociological review*, *72*(5), 784-811.

Banisch, S., Lima, R., & Araújo, T. (2012). Agent based models and opinion dynamics as Markov chains. *Social Networks*, *34*(4), 549-561.

Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A., & Bonneau, R. (2015). Tweeting From Left to Right Is Online Political Communication More Than an Echo Chamber? *Psychological science*, 0956797615594620.

Bednar, J., Bramson, A., Jones-Rooy, A., & Page, S. (2010). Emergent cultural signatures and persistent diversity: A model of conformity and consistency. *Rationality and Society*, *22*(4), 407-444.

Bimber, B. (1998). The Internet and political transformation: Populism, community, and accelerated pluralism. *Polity*, 133-160.

Boutyline, A., & Willer, R. (2013). *The social structure of political echo chambers: Ideology and political homophily in online communication networks*. Working paper, University of Berkeley.

Centola, D., & Macy, M. (2007). Complex contagions and the weakness of long ties1. *American journal of Sociology*, *113*(3), 702-734.

Conover, M. D., Gonçalves, B., Flammini, A., & Menczer, F. (2012). Partisan asymmetries in online political activity. *EPJ Data Science*, *1*(1), 1.

Dandekar, P., Goel, A., & Lee, D. T. (2013). Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, *110*(15), 5791-5796.

Bor, S.E. (2014). Echo Chamber Phenomenon. In *Encyclopedia of Social Media and Politics*. (pp. 427-428). Thousand Oaks. Sage Publishing.

Erisen, E., & Erisen, C. (2012). The effect of social networks on the quality of political thinking. *Political Psychology*, *33*(6), 839-865.

Feller, A., Kuhnert, M., Sprenger, T. O., & Welpe, I. M. (2011, July). Divided They Tweet: The Network Structure of Political Microbloggers and Discussion Topics. In *ICWSM.*

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the American public. *Annu. Rev. Polit. Sci.*, *11*, 563-588.

Flache, A., & Macy, M. W. (2011). Local convergence and global diversity from interpersonal to social influence. *Journal of Conflict Resolution*, *55*(6), 970-995.

Flaxman, S., Goel, S., & Rao, J. (2016). Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, nfw006.

Garrett, R.K., Gvirsman, S. D., Johnson, B. K., Tsfati, Y., Neo, R., & Dal, A. (2014). Implications of Pro- and Counterattitudinal Information Exposure for Affective Polarization. *Human Communication Research*, *40*(3), 309-332.

Gentzkow, M. & Shapiro, J.M. (2011). Ideological Segregation Online and Offline. The Quarterly Journal of Economics, 126(4), pp.1799–1839.

Goldenberg, J., Libai, B., & Muller, E. (2001). Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing letters,12*(3), 211-223.

Gruzd, A., & Roy, J. (2014). Investigating political polarization on Twitter: A Canadian perspective. *Policy & Internet*, *6*(1), 28-45.

Hanna, A., Sayre, B., Bode, L., Yang, J., & Shah, D. V. (2011). Mapping the Political Twitterverse: Candidates and Their Followers in the Midterms. In *ICWSM* (pp. 510-513).

Hart, W., Albarracín, D., Eagly, A. H., Brechan, I., Lindberg, M. J., & Merrill, L. (2009). Feeling validated versus being correct: a meta-analysis of selective exposure to information. *Psychological bulletin*, *135*(4), 555.

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter. *Journal of Computer-Mediated Communication*, *18*(2), 40–60.

Huckfeldt, R., & Sprague, J. (1987). Networks in context: The social flow of political information. *American Political Science Review*, *81*(04), 1197-1216.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology a social identity perspective on polarization. *Public opinion quarterly*, *76*(3), 405-431.

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1968). The peoples choice: how the voter makes up his mind in a presidential campaign.

Lazarsfeld, P., & Katz, E. (1955). Personal inf.luence. *Glencoe, Free Presse.*

Lerman, K., & Ghosh, R. (2010). Information contagion: An empirical study of the spread of news on Digg and Twitter social networks. *ICWSM*, *10*, 90-97.

Manin, B., Stein, E., & Mansbridge, J. (1987). On legitimacy and political deliberation. *Political theory*, *15*(3), 338-368.

Manjoo, F. (2011). *True enough: Learning to live in a post-fact society*. John Wiley & Sons.

Mill, J. S. (1869). *On liberty*. Longmans, Green, Reader, and Dyer.

Mutz, D. C. (2002). The consequences of cross-cutting networks for political participation. *American Journal of Political Science*, 838-855.

Mutz, D. C., & Mondak, J. J. (2006). The Workplace as a Context for Cross-Cutting Political Discourse. *Journal of politics*, *68*(1), 140-155.

Negroponte, N. (1996). *Being digital*. Vintage.

Parsons, B. M. (2010). Social networks and the affective impact of political disagreement. *Political Behavior*, *32*(2), 181-204.

Prior, M. (2005). News vs. entertainment: How increasing media choice widens gaps in political knowledge and turnout. *American Journal of Political Science*, *49*(3), 577-592.

Sears, D. O., & Freedman, J. L. (1967). Selective exposure to information: A critical review. *Public Opinion Quarterly*, *31*(2), 194-213.

Siegel, D. A. (2013). Social networks and the mass media. *American Political Science Review*, *107*(04), 786-805.

Stroud, N. J. (2007). Media Use and Political Predispositions: Revisiting the Concept of Selective Exposure. *Political Behavior*, *30*(3), 341–366.

Sunstein, C. R. (2009). *Republic.com 2.0*. Princeton University Press.

Yardi, S., & Boyd, D. (2010). Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of Science, Technology & Society*,*30*(5), 316-327.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *nature*, *393*(6684), 440-442.