The echo chamber effect in Twitter: does community polarization increase?

Siying Du and Steve Gregory

Abstract A recent article criticized social media platforms for failing to "mobilize society into action" long enough to address any major global issue. This is attributed to the simplistic design of current social media platforms, which encourage ideas to spread virally but do not support consensus formation which might lead to lasting social change. One reason for this could be the well known "echo chamber" phenomenon, whereby people tend to discuss issues only with other like-minded people. Social media has been blamed for encouraging the echo chamber effect and increasing polarization in society. For example, in Twitter, it is very common for users to be "followed" by others with similar views. Is this a reflection of real life or does Twitter actually increase polarization of views? This paper investigates this by comparing the Twitter *follows* network at two points in time and detecting communities in the network of reciprocated *follows* relationships. We find that new edges are (at least 3-4 times) more likely to be created inside existing communities than between communities, and existing edges are more likely to be removed if they are between communities. This leads to the conclusion that Twitter communities are indeed becoming more polarized as time passes.

1 Introduction

A recent article [1] highlighted the paradox that, although the use of social media has becoming increasingly widespread, it has not been able to "mobilize society into action" long enough to address any major global issue. The authors blame this on the simplistic design of current social media platforms, pointing out the absence of mechanisms for reflection, argumentation, and consensus formation. This is related to the well known "echo chamber" phenomenon, whereby people tend to discuss issues only with others with similar views. Social media has been blamed for encouraging the echo chamber effect and increasing polarization in society [2, 3].

Siying Du and Steve Gregory
Department of Computer Science, University of Bristol, Bristol BS8 1UB, UK

It is common knowledge that social networks, in real life as well as online, feature assortative mixing: people (or users) tend to communicate with those who are similar to themselves in some respect. When represented as networks, groups of vertices representing similar people tend to be more densely connected by edges than one would expect by chance [4]. This is the basis of community structure in networks, which has been studied intensively during the last 15 years [5].

In the context of online social media platforms, such as Facebook and Twitter, it is well known that user networks feature community structure. Users usually "follow" or "friend" other similar users, forming groups that are densely connected but loosely connected to other groups. When similarity is based on interests or opinions, users tend to be more strongly connected to others with similar interests and isolated from those with different interests or opposing viewpoints. One early study [6] analysed the network structure of (US domestic) political blogs and found that conservative and liberal blogs formed separate communities with little overlap. Following the launch of Twitter, another seminal work [7] obtained tweets related to a US election and constructed a *retweet* network, in which each edge represents a retweet from one user to another. This network was also found to split into two separate, ideologically opposed, communities.

The above works showed that social media platforms facilitate the echo chamber effect, by allowing users to form communities. However, this does not necessarily mean that these platforms encourage the formation of separate communities; they might have existed already.

The aim of this paper is to investigate whether social media platforms increase polarization of users, using Twitter as an example. We do this by checking whether community structure in the Twitter *follows* network becomes stronger, in some sense, as time passes. We consider only the network topology, ignoring the attributes of users and the content of their communication (tweets). We do not attempt to detect the topic or viewpoint that characterizes each community, or even verify whether a coherent topic exists. This is for simplicity and to avoid making our results dependent on a specific method of topic detection.

A naïve approach might be to perform community detection [5] on the network and compute the modularity [8] of the partition, and repeat the process at different times. However, this would be impractical because

1. The Twitter *follows* network is too large to obtain and analyse, especially because access to it is rate-limited.
2. The network vertices change over time as users come and go.
3. Different partitions could be found each time, as an artefact of the (nondeterministic) community detection algorithm.
4. Modularity (or some other common metric) depends on many factors and would not reveal small changes in the strength of community structure.

Our approach avoids these problems, as follows:

1. We collect small samples instead of the whole (reciprocated) *follows* network.

2. We sample the same set of users each time the experiment is repeated.
3. We detect communities only on the first run of the experiment.
4. We measure the strengthening of the community structure by counting how many new reciprocated *follows* edges are created inside communities and how many edges are removed between communities, and comparing these with a null model in which edges are added and deleted randomly.

In the next section, we explain how data is collected from the Twitter network. Section 3 presents the experimental results for new and deleted edges, comparing these with a randomly changed network. Section 4 presents our conclusions.

2 Data collection

The data collection was done in two phases: in June and August 2016. In each phase, three network samples were collected. This section describes the network samples and how they were collected.

2.1 First phase: snowball sampling

The basic strategy for the first phase of data collection is snowball sampling. This starts from a seed user (vertex) s and crawls to all of its followers (users who follow s) and "followings" (or "followees": users who are followed by s). This process is repeated recursively for each of the users found until enough vertices are obtained. We crawl to a maximum distance d from the seed, collecting all vertices at distance 0, 1, …, d-1, but not necessarily all vertices at distance d, because of the huge number of them.

In order to reduce the time costs, we choose a seed which has a reasonably small number of followers and followings. For our experiments we collected network samples from three seeds: a beauty blogger, a comic writer and a computer graphic scholar. We refer to these networks as "Beauty", "Comic", and "Graphics", respectively.

2.2 Omitting users or edges

Because of the rate limit of Twitter's API, which allows 15 requests every 15 minutes, it is time-consuming to collect users who have a large number of followings or followers. For example, if a user has 4 million followers, which is quite common for a famous person, it would take 13 hours to collect all of the user's followers. Because of the time cost and the limited time available, it was necessary to restrict the data collection.

One way to achieve this would be to omit users who have a large number of followers, and the other is to partially collect the followers and followings of a user. Both of these methods will introduce bias to the data collected. For the first method, we might miss a user who is famous and has an important role within a community (as well as all edges of this user). Although the first method is not perfect, the bias of the second method is much more severe. If we were to omit some edges between users, we are likely to miss some users who would form triangles with other users and create communities. For example, x, y, and z all follow each other, forming a triangle as shown in Fig. 1(a). If we partially collected followers of x and omitted z, which is a follower of x, the triangle $\{x, y, z\}$ might not be noticed, as in Fig. 1(b). As a result, the community detection might not place them into the same community, resulting in a distorted structure. Moreover, in this case, a deeper search might be needed to find z: in order to find z, one has to find y first. Obviously, the peripheral vertices will never be complete because the data collection has to stop somewhere, but we make sure that the network sample contains all edges for vertices that do appear in the sample. I.e., if the network sample is $G = (V, E)$ and $u \in V$ and $v \in V$ and $\{u, v\}$ exists in the complete network, then $\{u, v\} \in E$.

Therefore, we decided to omit all users who have more than 50,000 followers.

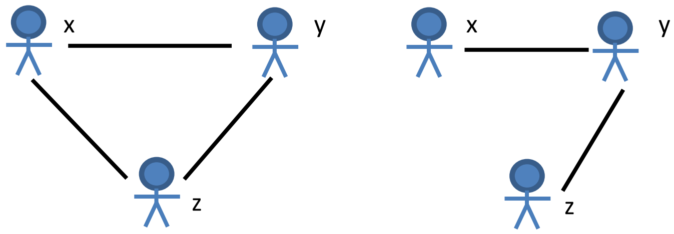

Fig. 1. (a) x, y, and z follow each other, forming a triangle. When collecting all followings and all followers of a user, this triangle can easily be found. (b) However, when collecting followings and followers partially, this triangle might be ignored.

2.3 Directionality

Considering the edge direction should be expected to contribute to a more accurate result [9]. However, in Twitter, any user u can follow any other user v, creating a directed edge (u, v). Such a unidirectional edge is less valuable than a reciprocated pair of edges, *u follows v* and *v follows u*, which indicate a mutual relationship. We therefore focus on undirected networks, in which an edge $\{x, y\}$ means that *x follows y* and *y follows x*. When collecting followers of a specific user u, we omit those users that u does not follow; when collecting u's followings, we omit users that do not follow u. To implement this, directed networks were collected and then converted to undirected networks with reciprocated edges after sampling.

2.4 Three datasets

In order to make our results more robust, we collected three different networks starting with three different seed vertices. Fig. 2 shows a visualization of the Graphics network, while Table 1 shows some statistics about all three networks collected in the first phase, in June 2016. This describes the three directed networks and the three undirected networks which contain reciprocated edges only. Table 1 also shows the communities found by the Infomap algorithm [10] for each network. We use Infomap for all experiments in this paper because it is one of the best and most popular community detection algorithms.

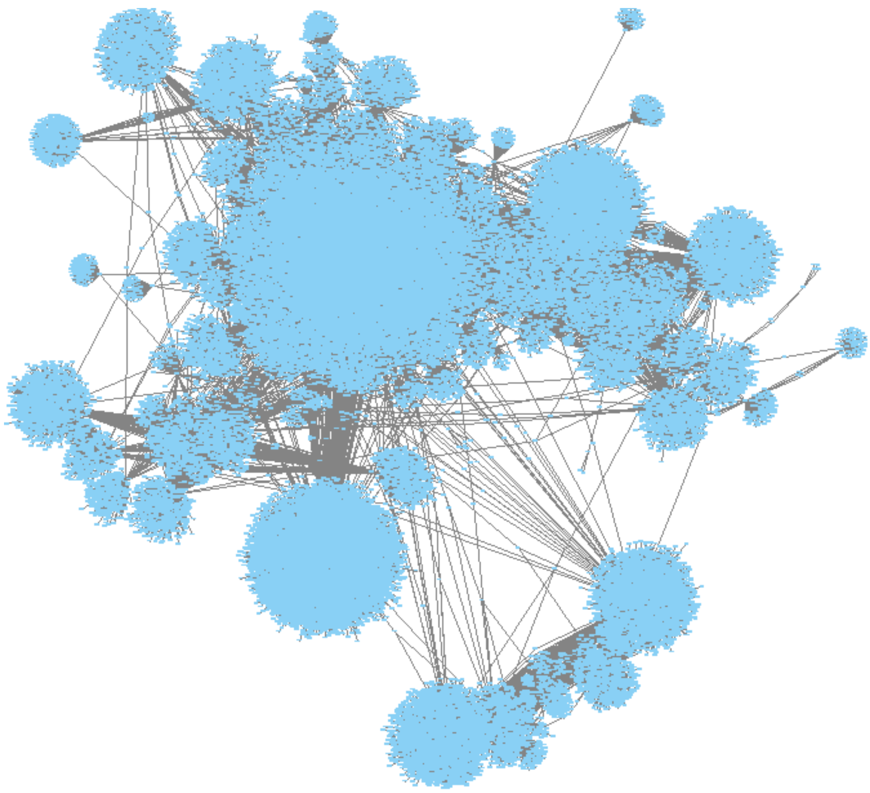

Fig. 2. Visualization of the Graphics network.

Table 1. Statistics of the three networks collected in June. The "Directed" columns indicate the vertices and edges before removing directionality. The "Undirected" columns describe the network of reciprocated edges.

Network	**Directed**		**Undirected**				
	vertices	edges	vertices	edges	density	communities	largest community
Beauty	6756319	10394337	249259	437852	$1.4\text{x}10^{-5}$	44	58275
Comic	2277503	3860175	101022	171990	$3.4\text{x}10^{-5}$	22	43681
Graphics	938960	1444554	47179	77909	$7.0\text{x}10^{-5}$	10	26563

2.5 Data collection for the second phase

There are two possible methods for the data collection of second phase (in August 2016). One is to crawl again from the same seed to collect a network by snowball sampling. The other is to directly collect all of the users that appeared in the first phase. Crawling from the beginning means doing a breadth-first search to a specific depth; this cannot ensure that all the users of first phase will be collected in the second phase. For instance, suppose that x, y, and z follow each other and form a triangle. When crawling from x with a depth of 1, this triangle will be found. However, if one of these edges is deleted before the second phase, a depth of 2 will be needed to find the triangle. As a result, crawling from the beginning with the same depth will omit some users that exist in the first phase, resulting in an incomplete network. Therefore, we chose to collect exactly the same users as in the first data collection phase, except those that no longer exist. Table 2 shows statistics about the same three network samples collected in August. (Note that, although we collect the same users as in the first phase, the number of vertices shown here is different because it includes all followers and followings.)

3 Experiments

3.1 Edges of real network

Community detection was performed on the network samples from the second phase, but did not show any noticeable changes because two months is not enough time for communities to evolve. However, there are still a significant number of new edges and deleted edges. The next step is to investigate how often new edges appear inside communities, indicating that users start to follow others in the same community, and whether edges tend to be removed (by "unfollowing") inside or between communities. We make two hypotheses:

Table 2. Statistics of the three networks collected in August.

Network	Directed		Undirected	
	vertices	edges	vertices	edges
Beauty	6957428	10717644	248363	463797
Comic	2546530	4214677	103353	180604
Graphics	994529	1522011	47491	84028

1. New edges are more likely to appear inside communities than between communities.
2. Edges between communities are more likely to be removed than those inside them.

In the remainder of the paper, we refer to edges inside communities as *intra-community* edges and edges between communities as *intercommunity* edges.

Fig. 3 shows the numbers of added and deleted edges of the three networks collected, counting only the edges between vertices that are present in both versions of the network. That is, we ignore vertices that existed only in the first snapshot, and their edges. For example, in the Beauty network, after two months, 5076 new edges appear: 3212 intracommunity edges and 1864 intercommunity edges. Similarly, in the other two networks, most of the new edges are intracommunity edges, which seems to support the first hypothesis stated above. For the deleted edges, for all three networks, the number of intracommunity deleted edges exceeds the number of intercommunity deleted edges, which seems to disprove our second hypothesis. However, intracommunity edges are far more numerous than intercommunity edges, so whenever an edge is removed, it is more likely to be an intra-community edge, by chance.

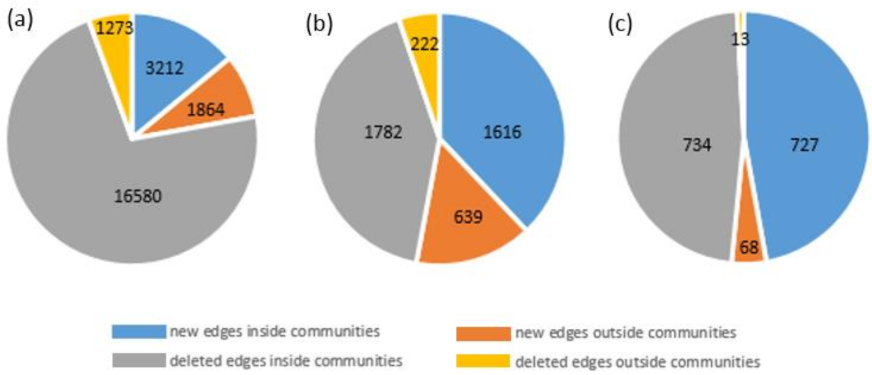

Fig. 3. Distribution of new and deleted edges of the three networks collected in both phases. (a) Beauty; (b) Comic; (c) Graphics.

3.2 New edges of random case

To evaluate the numbers of new and deleted edges correctly, the actual numbers must be compared with a null model which adds or deletes edges randomly.

If $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are the networks of the first and second phase respectively, we randomly generate a new edge $\{u, v\}$ where $u \in V_1 \cap V_2$, $v \in V_1 \cap V_2$, and $\{u, v\} \notin E_1$. This means that we connect a randomly chosen pair of vertices that existed in both June and August but were not linked by an edge in June.

Based on this strategy, for every network, the total number of edges added is equal to the number in the corresponding real network. Table 3 shows the average number, largest number, and smallest number of new *intracommunity* edges in all three networks after generating the randomly grown network 100 times. Taking the Graphics network as an example, there should be 795 new edges, of which 727 are intracommunity (calculated from Fig. 3). From this table, the average number of intracommunity new edges in the random case is 297 which is much less than the real result, which is 727 edges. Even the largest value found, 329, is still much less than 727. For the other two networks, the results are consistent with the Graphics network. This answers our question: new edges occur inside communities more often than expected by chance.

Table 3. Intracommunity new edges of the random case in the three networks.

Network	Maximum	Minimum	Average	Real
Beauty	704	584	**637**	**3212**
Comic	652	548	**597**	**1616**
Graphics	329	254	**297**	**727**

Fig. 4 shows the distribution of the number of intracommunity edges added in each of the random networks. The star in each chart represents the number of intracommunity edges in the corresponding real network, which is always much greater than the numbers achieved in the random case. This allows us to reject the null hypothesis that the result is by chance. In principle, we could plot these curves analytically and calculate the extremely small probability that the real result could happen by chance, but we have not done so here.

3.3 Deleted edges of random case

Fig. 3 shows that most of the deleted edges are intracommunity edges, but this is to be expected because there are relatively few intercommunity edges to delete. We need to investigate whether deleted edges are more likely to be intercommunity than expected. We do this by simulating another shrunk network based on the original network. The strategy is to remove edges from this network randomly.

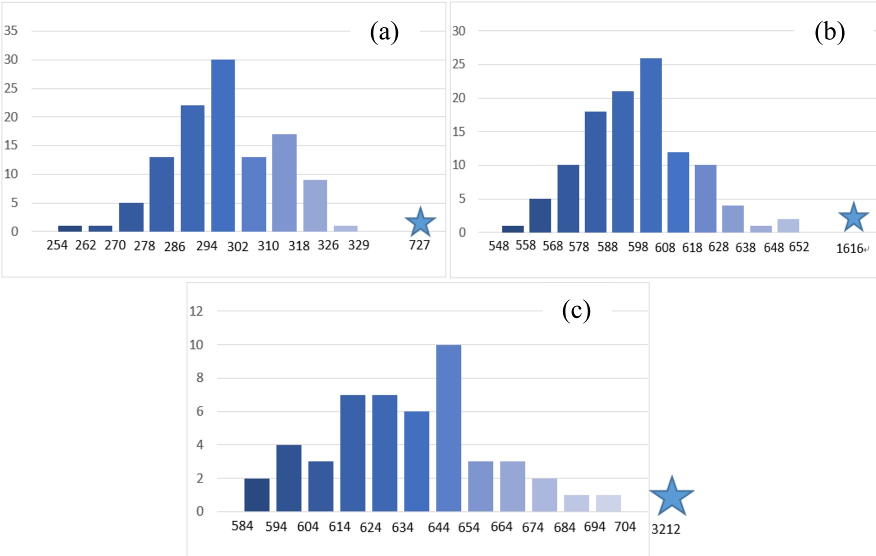

Fig. 4. Number of new intracommunity edges added in each network. (a) Graphics network (run 112 times); (b) Comic network (run 110 times); (c) Beauty network (run 50 times).

If networks $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are the networks of the first phase and second phase respectively, we randomly choose a edge $\{u, v\}$ where $u \in V_1 \cap V_2$, $v \in V_1 \cap V_2$, and $\{u, v\} \in E_1$. This means that we randomly choose a pair of vertices that existed in both June and August and were linked by an edge in June, and delete that edge.

Fig. 3 shows the number of deleted edges in the three networks. For these three networks, 17853, 2004, and 747 edges were removed, respectively.

In order to test the hypothesis that intercommunity edges are more likely to be deleted, we compare the number of deleted edges of the random case with the real network, in Fig. 5. Taking the Comic network (Fig. 5(b)) as an example, the average number of intercommunity deleted edges is around 75 and even the maximum, 97, is far less than the real result, 222. These results are less pronounced than for added edges (Section 3.2) but still show that intercommunity edge deletion is more common than expected by chance.

3.4 Biased network

Section 3.2 showed that new intracommunity edges are added far more often than could happen by chance, but a more interesting question is *how much* more often.

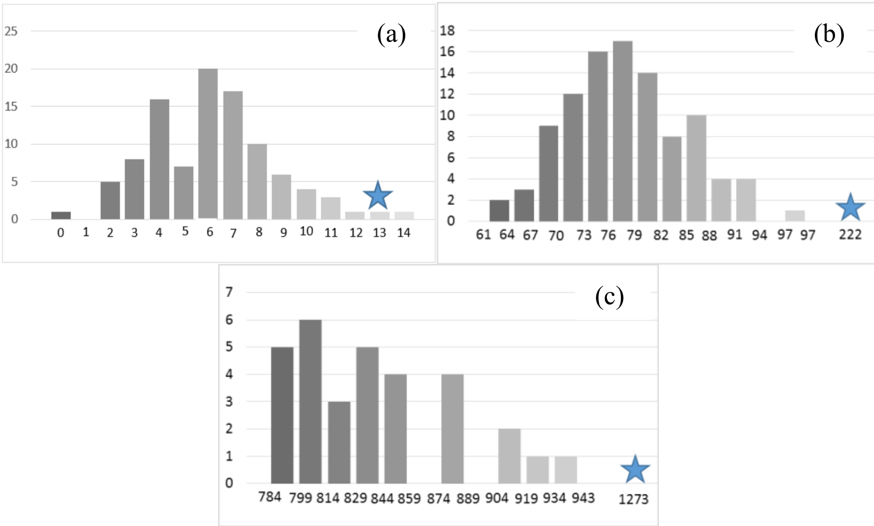

Fig. 5. Number of deleted intercommunity edges in each network. (a) Graphics network (run 100 times); (b) Comic network (run 100 times); (b) Beauty network (run 31 times).

In order to measure this, we imagine a biased random agent that repeatedly adds new edges: each edge has a probability p to be an intracommunity edge; otherwise it is an intercommunity edge. We adjust the probability p until the number of intracommunity edges added is close to the real value. After testing several times, the probability values found for the three networks are 0.7 (Beauty), 0.75 (Comic), and 0.82 (Graphics). From Table 3 and Fig. 3, we can compute equivalent probabilities for an unbiased random agent: 0.12, 0.26, and 0.37, respectively. This means that, in the Beauty network for example, intracommunity edges are nearly six times more likely to be added than expected by chance.

4 Conclusions

We have shown that, at least for three network samples, the community structure of the Twitter "follows" network seems to become stronger as time passes, increasing the separation between communities.

It is important to emphasize that we have analysed only the network topology and not the details of the users or their tweets, which are outside the scope of this work. Therefore, we have no evidence of whether a community (in our sense) represents a single topic or viewpoint, or whether different communities represent opposing viewpoints. Indeed, because we only detect *disjoint* communities, it is unlikely that each community detected discusses only a single topic. Nevertheless, in cases where communities do correspond to viewpoints, this separation can be interpreted as polarization.

Our specific findings are:

1. New edges are intracommunity edges much more often than expected.
2. Deleted edges are intercommunity edges much more than expected.
3. When adding edges, users are about 3-4 times more likely to add an intra-community edge than an intercommunity edge.

These observations probably underestimate the true effect. Because we collect small samples of the network, community detection is certain to be imperfect because some communities are split between the sample and the rest of the network and cannot be found. In the extreme case, if random communities were found, our results would be no different from the random null model with which we compare. If we had time to collect larger samples, we would therefore expect an even more pronounced effect. This is a good topic for future work.

It is interesting to speculate on the reason for the effect we observe. One possible explanation is the recommender system of Twitter: users receive suggestions about users that they might want to follow, and these are often users who are already in the same network community. Further work would be needed to find out whether the generation of new edges is consistent with Twitter's recommendations (which are not revealed except to the users themselves). In any case, the recommender system cannot be the only explanation because of (2) above: Twitter never recommends users to "unfollow". It seems more likely that users start following others after discovering them through the network structure itself; e.g., by retweets. New users (those that exist in the later snapshot but not the first) might even play a role in introducing existing users to each other and causing an edge to appear, even though we exclude these new users from our network samples.

Finally, it may be argued that, even if Twitter communities become more polarized over time, this might not be caused by the platform itself. The Twitter network may be converging over time to an underlying real-world network which is already highly polarized. Even so, Twitter provides the mechanisms to reflect and enhance this polarization, unlike traditional media and communication methods, which might tend to reduce it.

Future work

Further work is needed to estimate the probability with which a biased random agent chooses an intercommunity edge to delete.

We have used a simple null model for our unbiased random agent, whereby vertices to connect are chosen uniformly randomly from all vertices in the sample. Numerous other null models are possible; for example, the agent might preferentially connect to popular (high-degree) users or to users with a similar name or description. In future, it would be useful to test other null models to rule out other possible explanations for the results found.

Another area of future work is to repeat the analysis with different community detection algorithms instead of Infomap. This is simple to do because we have kept the sampling and analysis phases separate, which would not be the case if we had used (e.g.) a local modularity [11] method to collect the network samples. A more challenging task would be to detect overlapping communities, instead of disjoint communities, in the networks. Overlapping communities are more realistic because many Twitter users have more than one interest and hence belong to multiple communities. However, overlapping community detection is more difficult and the results would be harder to analyse.

Acknowledgements

We are grateful to the anonymous referees for insightful comments that have improved the final version of the paper, especially its conclusions.

References

1. Cebrian M, Rahwan I, Pentland A (2016) Beyond viral. Commun. ACM 59(4):36-39
2. Dewey C (2013) How Twitter makes the political echo chamber worse. Washington Post, 3 Sep 2013. https://www.washingtonpost.com/news/the-fix/wp/2013/09/03/how-twitter-makes-the-political-echo-chamber-worse/ Accessed 18 Sep 2016
3. DiFonzo N (2011) The Echo-Chamber Effect. New York Times, 22 Apr 2011. http://www.nytimes.com/roomfordebate/2011/04/21/barack-obama-and-the-psychology-of-the-birther-myth/the-echo-chamber-effect Accessed 18 Sep 2016
4. Newman M (2003) Mixing patterns in networks. Phys. Rev. E 67:026126
5. Fortunato S (2010) Community detection in graphs. Phys. Rep. 486:75-174
6. Adamic L A, Glance N (2005) The political blogosphere and the 2004 U.S. election: divided they blog. In Proc. of the 3rd Intl. Workshop on Link Discovery. ACM, 36-43
7. Conover M D, Ratkiewicz J, Francisco M, Goncalves B, Flammini A, Menczer F (2011) Political polarization on Twitter. In: Proc. of the 5th Intl. AAAI Conf. on Weblogs and Social Media
8. Newman M, Girvan M (2004) Finding and evaluating community structure in networks. Phys. Rev. E 69:026113
9. Amor B, Vuik S, Callahan R, Darzi A, Yaliraki S N, Barahona M (2016) Community detection and role identification in directed networks: understanding the Twitter network of the care.data debate. In: Adams N, Heard, N (eds) Dynamic networks and cyber-security. World Scientific Press
10. Rosvall M, Bergstrom C T (2008) Maps of random walks on complex networks reveal community structure. Proc. Natl. Acad. Sci. USA 105(4):1118-1123
11. Clauset A (2005) Finding local community structure in networks. Phys. Rev. E 72:026132