



Tweeting Russian Politics: Studying Online Political Dynamics

Mikhail Zherebtsov and Sergei Goussev

30.1 INTRODUCTION

The popularity of social media studies in the context of Russian politics started to take off during the 2011–2012 civil protests against what are widely seen as fraudulent results of the 2011 parliamentary elections. In the absence of impartial and objective coverage of elections in the traditional media (Golos 2011) various Social Network Sites (SNS) appeared to be instrumental in circulating information among citizens, effectively serving as an alternative source of trustworthy information on the election process. The key feature of social media functionality during the 2011–2012 protest was its multi-channel and multi-hierarchical structure. Spontaneously emerging information posted online about fraud and other infringements over the course of the elections was picked up and popularized by famous bloggers and popular public channels. It helped build awareness of the magnitude of committed acts and reaffirm large public discontent regarding the validity of the election process and the pronounced winner—the pro-Kremlin *Edinaá Rossiá* (United Russia) party. Furthermore, social media were instrumental in organizing and coordinating the protest as a key means of information circulation (for more on digital activism, see Chap. 8). Earlier academic accounts were positive on the crucial role of SNS in

M. Zherebtsov (✉)
Institute of European, Russian and Eurasian Studies, Carleton University,
Ottawa, ON, Canada
e-mail: mikhail.zherebtsov@carleton.ca

S. Goussev
Ottawa, ON, Canada

“mobilizing the discontent of citizens under the conditions of a semi-authoritarian political regime” (Lonkila 2012, 9).

This type of research received impetus from a chain of protests, ranging from the Occupy movement in mostly developed countries to the Arab Spring in the Middle East and North Africa region. In these cases, the internet in general and social media in particular were the key factors in the scope and magnitude of the protests. The events spurred a vigorous research into the phenomenon of social media in the context of social capital and civic engagement (Agarwal et al. 2014; Fuchs 2014), mobilization of protests (Earl et al. 2013; Greene 2013; Breuer et al. 2015), as well as methodological boundaries of political science, associated with new and growing computational methods (Tremayne 2014; Sinclair 2016; Tucker et al. 2016).

The events of the Arab Spring further reinvigorated the discourse of democratization of authoritarian states, claiming social media to be the key underlying technology promoting political change (Tufekci and Wilson 2012). These studies have fallen on the fertile soil of the Russian protest realities, determining the key theme of research. Therefore, in the context of politics, social media and digital social networks, Russian studies explored the main democratization hypothesis (Greene 2013), looking for answers as to why the Russian case did not result in a tumultuous upheaval akin to the Arab Spring (White and McAllister 2014; Reuter and Szakonyi 2013; Pallin 2017). In their answers, authors outlined several key features of social media in Russia. First of all, they agreed that up until the 2011–2012 protests the Russian digital public sphere had been developing relatively freely, without the tight oversight of the government. Secondly, recognizing the importance of Internet technologies, the authorities preferred a rather flexible model of domination over the rigid regulatory framework, which are popular in other authoritarian countries (with China being the exemplary case). On the one hand, an active and popular pro-government audience was cultivated and exhibited to the entire political spectrum, on the other, any anti-government sentiment was disrupted by various means, including the use of bots and trolls. Finally, the government undertook measures to domesticate the ownership over key social network sites and to ensure compliance of the large international ones through excessive regulation.

Following the emergence and controversy around the cyber activities of Russian government-affiliated organizations outside Russia’s territorial border, the research agenda and discourse then shifted towards a deeper study of trolls and bots (Jensen 2018; Stukal et al. 2017). Therefore, given the public interest in these specific topics, Russian social media studies have been dominated by a rather narrow research agenda, mainly referring to abnormal and critical situations. Obviously, patterns of users’ behavior would differ during these events, from their behavior under normal circumstances.

There have been earlier attempts to depict the topology of Russian digital social networks (Barash and Kelly 2012; Kelly et al. 2012) as well as the contents of key discussions (Nagornyy and Koltsova 2017; Maslinsky et al. 2013), using advanced computational methods of Social Network Analysis (SNA) and

topic modeling. Yet, apart from specific and quite narrow-focused contributions from other fields and especially computational linguistics, such studies remain on the periphery of contemporary policy research.

30.2 GOALS AND DATA

This chapter aims to fill this gap by analyzing the intra- and intergroup structure of politically engaged users of Russian social media and presenting the dynamics of ongoing political discussions. It builds on already conducted research and applies key SNA methods and approaches to the corpora of social media data from the Russian segment of Twitter. Therefore, the goals of the chapter are twofold: (1) to demonstrate the potential of SNA to analyze political discussions in Russian social media, and (2) to establish online political communities, determine their internal structure as well as measure their interconnectedness and detect key influencers. The chapter explores a number of hypotheses regarding the role of social media in contemporary Russian politics, which were partially inspired by the previous research in the field.

We propose and test a three-tier analytical strategy, outlining the macro-, meso-, and micro-levels of network analysis. We check whether Russian virtual society is generally divided across the same ideological lines as the public sphere, representing two scattered, yet distinct groups representing the pro-government sphere on the one hand, and, mainly, “non-systemic” opposition forces on the other (Gel'man 2015). Applying automated community-detection methods, we determine and visualize existing online communities. We further analyze their characteristics and, comparing their user structure with the contents of selected discussions, determine their ideological basis. Finally, we investigate relationships between and within communities, establish key leaders and influencers, as well as test the possibility of a dialogue between existing online political communities (for another case of network analysis, see Chap. 29).

While other social media platforms are more utilized by the wider public in Russia, such as VK (formerly VKontakte) or OK (formerly Odnoklassniki), we focus on Twitter due to three factors. Firstly, as a micro-blogging platform, it is determined by its inherently public nature, allowing the sharing and viewing of content without a restrictive permission structure, even permitting the viewing and following of all public content without a Twitter account. In part due to this, for public figures the platform has become a sort of *modus operandi*, both amongst the pro-government and “non-systemic” opposition. In this regard, Twitter, to a certain extent, has complemented LiveJournal—another very popular blogging platform as a means of reaching to a wider audience, yet with shorter announcements and “punchier” messages. Twitter remains the sixth most important social network platform in Russia with 9.9 million of unique monthly users.¹ In absolute numbers, Russian Twitter segment is the fifth largest in the world in terms of active user accounts.² Moreover, the Russian segment appears to be the most politicized, as a high proportion of the

top 100 most followed accounts are political figures and media accounts, compared to other countries.³ Secondly, Twitter has been a platform highly utilized for political information dissemination and event coordination, including for protests, both internationally, such as during the Arab spring, and in Russia, particularly during the 2011–2012 protests (Lotan et al. 2011; Wolfsfeld et al. 2013; Spaiser et al. 2017). Thirdly, it has been argued that foreign-owned social media (Facebook and Twitter) have a greater impact on the patterns of circulation of anti-government and pro-protest information than domestically owned platforms (VK, OK), due to greater state control over the latter ones (Reuter and Szakonyi 2015). Taken together, these factors demonstrate that Twitter in Russia is an important and contested space for the politically engaged segment of the Russian population and is a relevant and valuable platform for the political analysis of the country. Therefore, given its inherently open and public nature and high nominal politicization, Twitter in Russia is regarded as a valuable object for analysis of politically active social networks and political communication in the country.

To perform the empirical assessment, we collected six samples of Twitter data on topics of international or national political importance. Given their extensive coverage in traditional media altogether with higher than average Twitter activity, each event demonstrated resonance in Russian political society (see Table 30.1 for list). Each sample was collected individually using Twitter's Search feature of the REST API, which allows the retroactive extraction of recent popular tweets containing specific keywords and returning a sample of tweets made in the preceding 7 to 9 days. The advantage of this approach is that it allows the collection of content preceding, during and following each specific informational event. To construct and evaluate user communities for each event, which are commonly understood to be based on who each user chooses to follow (Colleoni et al. 2014; Barberá et al. 2015; Halberstam and Knight 2015), we collected data on all friends/followers of users who participated in the sampled political discussions. This approach results in the collection of event data, participant users, and relationships between them. The total corpora of all six samples included 175K users and 978K tweets and retweets.

30.3 ASSESSMENT OF SNA METHODS

30.3.1 *Macroscopic Methods: Visualizing Russian Online Political Communities*

The common starting point of network analysis is the detection of network structure and visualization of the resultant communities. Visualization allows at-a-glance assessment of the patterns present in between captured entities and the identification of which subsequent methods are relevant to investigate specific details of relationships of interest. Among various graph-visualization methods, force-directed layouts have become highly popular for practitioners in part due to the fact that they are aesthetically pleasant and intuitive (Koren

Table 30.1 Six events and size of captured networks

<i>Event</i>	<i>Date of event</i>	<i>Timeline captured</i>	<i>Total dataset captured</i>	<i>Keywords used (translation)^a</i>
Two-year anniversary (2016) of the accession of Crimea (collecting pro- and anti-Russian sentiment towards the event)	Mar 18	Mar 15– Mar 25	Users (nodes): 7,772 Tweets: 12,919 Edges: 397K	#Krymnaš (#CrimeaIsOurs)
Eurovision Song Contest 2016; the victory of Ukrainian singer Džamala	May 14	May 8–May 22	Users: 118,725 Tweets: 364,734 Edges: 8.49M	Džamala (Jamala); Evroviedenis (Eurovision)
Dmitrij Medvedev’s public comment to pensioners “there is no money, but you hang in there”	May 23	May 20– May 30	Users: 17,932 Tweets: 44,324 Edges: 691K	Medvedev (Medvedev)
Release of Ukraine’s prisoner of war Nadežda Savčenko in a prisoner swap	May 25	May 21– May 30	Users: 42,625 Tweets: 200,350 Edges: 3.17M	Savčenko (Savchenko)
Liberal governor Nikita Belyh’s arrest for corruption	Jun 24	Jun 19–Jun 28	Users: 25,272 Tweets: 84,992 Edges: 1.56M	Belyh (Belykh)
Turkey’s 2016 failed military coup against president Erdogan (responses of Russian audience, following Russian-Turkish discord resulting from the shooting down Russian fighter jet in 2015)	Jul 15	Jul 16–Jul 19	Users: 44,947 Tweets: 271,391 Edges: 3.60M	Turci* (Turkey); Tureck* (Turkish); Erdogan* (Erdogan); vosstani* (rebellion); perevorot* (coup)

^aKeywords and hashtags were used to collect a focused discussion sample and to minimize unrelated discussions. For keywords specified with a *, all possible keyword inflections were utilized in the search

2003). Force-directed layouts, such as the commonly utilized ForceAtlas2, simulate a natural physical system of forces acting upon each other, with nodes repulsing each other like charged particles and edges attracting nodes like springs (Bastian et al. 2009; Jacomy et al. 2014). Applied to social networks such as the Twitter follower network, the method visually clusters

well-connected users and segregates loosely affiliated groups. It furthermore maps the network, practically visualizing the distance between users and user groups.

Modularity and community-detection methods quantify key network parameters, complementing visualization layouts. The modularity statistic measures how divided a network is into segregated groups, ranging from -0.5 to 1 , with the upper range indicating stronger module segregation (Brandes et al. 2007). Community-detection algorithms analyze the network structure and assign nodes to communities, providing a statistical analysis that complements the visual representation of force-directed graphs. The established network parameters support the evaluation of sociological theories and research hypotheses, such as the presence of “echo chambers” in social networks. Based on the hypothesis that most engagement happens amongst likeminded and connected individuals (Bakshy et al. 2015; Colleoni et al. 2014), this phenomenon has been widely investigated in international cases (Colleoni et al. 2014; Barberá et al. 2015); however it is hitherto under researched for the Russian case.

Choosing an appropriate community-detection method depends on the network type as well as computational resources, with particularly large-sized networks presenting a challenge. For directed networks, as in this case, with edges representing follower relationships or communication patterns like retweets or mentions, the Infomap method is appropriate (Lancichinetti and Fortunato 2009).⁴ The Infomap method (Rosvall and Bergstrom 2008) simulates a random walk along the edges of the network and categorizes communities where information can flow quickly amongst well-connected users and is unlikely to leave to another group (Rosvall et al. 2009). We apply the Infomap method to categorize user communities on the six captured political samples, calculate modularity, and visualize each using the ForceAtlas 2 (Jacomy et al. 2014) force-directed layout (see Fig. 30.1).

We observe that the political Russian Twitter space contains a highly stable community structure that parallels the real political landscape in the country, with two major political communities and a multitude of smaller ones, reacting to all political events in the country. The collected data allows us to assess various characteristics of established communities. Particularly, the basic follower method is complemented with an evaluation of network structures based on typical Twitter activities such as retweets and mentions. Both are useful as a retweet can be a symbolic representation of the consonance of opinions or importance of specific information, whereas mentions provide a wider spectrum of reactions and relationships between users. As hypothesized, there is division into two major competing political forces (Gel'man 2015), with the two major communities being (1) the pro-Kremlin (pro-government) supporters and Russian nationalists (community 0 or purple), and (2) the liberal and non-systemic opposition (community 2 or teal).

We also found that the “echo chamber” theory applies well to the Russian Twitter network, as the follower-based communities were highly polarized.

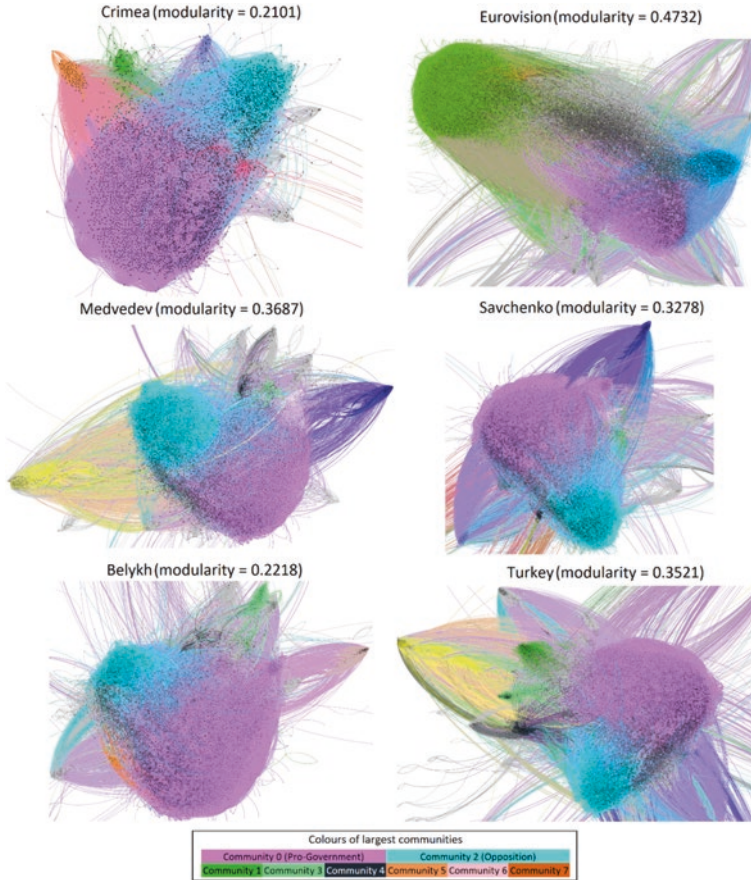


Fig. 30.1 The structure of political communities on Twitter by event

Modularity varied between the events, from a relatively low statistic of 0.2101 on the Crimea sample to a moderately strong statistic of 0.4732 for the Eurovision sample (see Fig. 30.1).⁵ Furthermore, users in each community showed a strong preference to retweet, mention, and communicate with users in their own community and low preference to do the same for users in other communities. As Tables 30.8, 30.9, and 30.10 (Annex A) show, on average 75% of all mentions, retweets, and replies happened within, and only 25% happened between users of different groups. Specifically looking at the pro-government and Opposition communities, we see that they are very highly polarized, as they retweet on average only 5% of the content created in the rival group. Interestingly, the Opposition group is less polarized of the two, possibly due to being on average three and half times smaller than the pro-government group.

30.3.2 *Mesoscopic Methods and Russian Political Communities: Similar or Different?*

Upon detecting and visualizing a macro structure of the whole network, it is useful to detail each of the detected communities through the two methods of density and transitivity. Both demonstrate the compactness of each community network, showing whether a group's users are only loosely connected or highly interrelated and hence likely to be ideologically contiguous. Whereas density approaches the network holistically, measuring the proportion of connections that are present in the network against the total number of possible connections, transitivity measures the proportion of triangles (or three users connected to each other) against all possible triangles, a stronger indicator of interrelationships. Therefore, networks that have high density but low transitivity will be relatively interlinked, but not all users will know each other.⁶ In practical terms, naturally built tight communities signify the presence of numerous multi-user interactions and the sharing of social trust and social capital within the group (Coleman 1990).

A further method is to detect cliques in a network, or a subset of nodes that are all connected to all other nodes in the clique. In social networks, cliques are sometimes referred to as clique communities, where groups of users are completely interconnected, with larger communities often containing many cliques. A benefit of clique analysis is that the prevalence and average size of cliques in a community network provides insight into the structure of the political group. A community with a large group of tri-node cliques (triads) demonstrates a relatively dispersed community, whereas the presence of several cliques with a large number of nodes in each hint at relevant sub-community structure for further analysis. Furthermore, as information is disseminated on Twitter through follower relationships, cliques represent a method of evaluating information propagation through a community, as well as a detailed analysis of the behavior of users in one versus another clique, as individuals tend to be highly influenced by the clique they belong to (Borgatti et al. 2009).

We find that the identified main political communities in Russian Twitter have vastly different characteristics and vary by event. The opposition community is a relatively dense and closely knit group, generally having stronger ties between individuals and likely sharing more meaningful interpersonal relationships. The pro-government community on the other hand is a more loosely related group of independent mini-communities, possessing more sporadic links between the sub-groups. In all six samples, the density of the opposition community exceeded that of the pro-government group (Table 30.2). Looking at transitivity, the pattern is repeated, although not as strongly and not for every sample. Clique distribution further underlies the social structure of both, as cliques in the opposition tend to be much smaller (Figs. 30.2 and 30.3). The looser amalgam of large cliques in the pro-government group also underlies the importance of public opinion leaders to reach each of these larger mini-communities.

Table 30.2 Density and transitivity of the network in its entirety and within its main communities

<i>Event</i>	<i>Density</i>			<i>Transitivity</i>		
	<i>Network</i>	<i>Pro-govt community</i>	<i>Opposition community</i>	<i>Network</i>	<i>Pro-govt community</i>	<i>Opposition community</i>
Belyh	0.41	0.92	2.64	9.85	10.73	18.12
Crimea	0.88	1.96	2.62	11.82	12.69	18.29
Eurovision	0.10	0.53	1.13	11.49	8.24	10.72
Medvedev	0.42	1.20	1.83	13.61	16.74	13.52
Savchenko	0.30	0.82	1.22	9.28	10.45	10.95
Turkey	0.23	0.70	1.48	9.20	9.49	11.80

Note: Displayed as percentage due to small scale

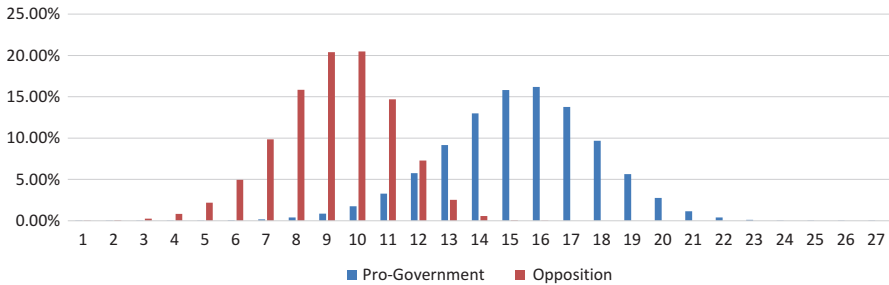


Fig. 30.2 Clique size frequency distribution by community—Crimea sample

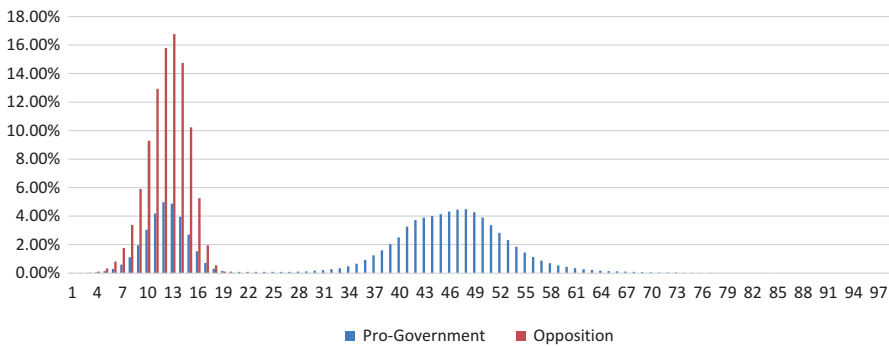


Fig. 30.3 Clique size frequency distribution by community—Medvedev sample

An important remark concerning the utilization of meso methods that is applicable to both Russian and international contexts, has to be made, however. Both, density and transitivity could be sensitive to the quality of sample data. For instance, keyword or hashtag searches could miss statements and,

hence users, that indirectly reference the political event. This would inevitably affect the subnetwork structure. Furthermore, the rate limits and index algorithms, used by Social Media Platform APIs, could also seriously impact meso methodologies (Pfeffer et al. 2018). One way to alleviate sample issues is to use multi-sample approaches to demonstrate cross-event consistency, as done in this study. Another is sampling using only general limitations, such as language or location. While Twitter's free API does not offer location filtering, language filtering has a potential for Russian political analysis, as fewer (compared to international languages) users outside the country would engage in online political discussion.⁷

30.3.3 *Microscopic Methods: Opinion Leaders in Russian Online Political Networks*

Following the evaluation of network structure and community sub-structure, scholars often turn to the identification and measurement of the impact of network's "influencers," as well as the comparison of these influencers to offline opinion "leaders." Traditional elites, who have always had the ability to shape the political narrative, have seen their power greatly expanded with Twitter and other social media spaces (Jungherr 2014). Previous research, both internationally (Bakshy et al. 2011) and on Russia (Roesen and Zvereva 2014) has found that traditional "leaders" can be cumulatively overshadowed on social media by "ordinary influencers" (Bakshy et al. 2011, 8), or median public figures with an average "offline" influence.

30.3.3.1 *Identifying and Evaluating "Influencers"*

The analysis is based on a sample of 469 accounts, which comprises public personalities and organizations as well as traditional media. These users were selected if they: (1) actively post on politically relevant events; (2) have at least ten thousand followers; and (3) either occupy positions in a government/non-government organization, or are well-known media personalities. The sampling technique adapted the "snowballing" approach but required several stages in order to improve the validity of the outcome. First, a top tier of politically relevant users was manually selected from the list of top 100 most popular accounts in the Russian segment. Secondly, from all samples collected, the 1000 most followed accounts were selected and manually sorted in order to identify politically relevant ones. These two steps together resulted in a list of 240 accounts. Among these accounts, only those that followed no more than 500 others, were selected. Subsequently, the list of friends of each was obtained, but only those who themselves had at least 10,000 followers were selected. Qualitative filtering of this list resulted in the creation of the master sample of 469 active Twitter public personalities.

The selection process inherent with this type of sampling technique can be considered as establishing a representative collection of user accounts. Given the "echo chamber" effect, it can be assumed that those who use Twitter as an

interactive platform, and not only spread but also receive information, will strategically connect with (or themselves follow) a limited amount of personalities, many themselves public figures and involved in analogous activities (politics in the case of this study). Although the 10,000-follower threshold is rather arbitrary, it allows the selection of only those accounts that have the potential to efficiently create and/or disseminate political information. Similarly, the 500-friend threshold excludes those personalities who apply a tactic of following any account that interacts with them, and whose inclusion would not improve the sample.⁸ To support the assessment of the endurance and impact of content created by opinion leaders, the last 3200 tweets of each were downloaded using the REST API and ranked in terms of their impact on political discussions. Four types of politically relevant leader accounts can be identified in the Russian Twitter segment. The first type are personal accounts of top politicians, media, and public personalities. Many of these accounts can be regarded as official, as they are verified “*de-jure*,” while others produce content that corresponds with ideological views of their nominal owners and therefore can be regarded as “*de-facto*” genuine. The second cluster comprises of accounts of traditional media sources, which utilize the platform predominantly to reach a wider user audience. In most instances, tweets produced by these types of accounts contain links to materials issued on these media’s websites, sometimes with opinionated comments that reflect the editors’ ideological preferences. These accounts appear to be the most interconnected within as well as outside the ideologically bounded communities they belong to. The third type includes official accounts of government agencies, which were selected for analysis on the basis of multiple premises. Twitter has been actively used by private sector companies and entrepreneurs for marketing purposes. Indeed, there is a growing body of research on the subject matter, which explores and analyses strategies of efficient public relations and marketing for businesses. If used efficiently, Twitter could boost a company’s performance. The same logic is applied to political organizations (Waters and Williams 2011; Towner and Dulio 2012), who adopt advanced technologies of governance within the Government 2.0 paradigm. This approach was officially adopted in Russia in the context of the Federal Program “Information Society 2011–2020” (Zherebtsov 2019). Accounts that produce and circulate political satire and politically relevant entertainment content comprise the fourth type of accounts, which we conventionally refer to as the parody group. While they themselves are not sources of official information or representatives of certain political groups, such accounts appear at the epicenter of selected discussions and disseminate certain sentiments. Moreover, they are quite popular not only among regular users but also among top political influencers.

The analysis of content produced by the leaders reveals several remarkable trends. There is a certain consistency between the groups in terms of retweeting and liking messages. The parody group outperforms all others in the combined popularity of its messages. Needless to say, all accounts in our sample that belong to this group produce and share oppositional sentiment. Personal

accounts of political leaders comprise the second most popular group on Twitter. Interestingly, the content produced by these types of accounts is as often retweeted (or shared and thus actively endorsed) as it is liked (or passively endorsed). The Twitter activity of traditional media appears to be much lower than the first two types of accounts. To some extent, this demonstrates a quite remarkable pattern of social engagement in the Russian Twitter segment. While entertainment purposes are prevalent even in the context of political discourse (as demonstrated by the overwhelming popularity of parody accounts), users tend to get involved in political discussions and favor opinionated statements of political pundits and media personalities over factual information circulation. This being said, it was to be expected that official accounts appeared to be the least publicized in our sample; a trend best explained by the nature of content produced and shared by the accounts of this group. As official accounts tended to share links to digests and press releases, produced by the press-services of their respective agencies, this information is regarded as the least entertaining (or “infotaining”) to users.

Table 30.3 illustrates the ranking of leaders’ accounts by popularity in terms of both active (retweets) and passive (likes) endorsements, both on average for all accounts over the entire sample collected, and using a subsample of the top 10,000 most popular tweets authored by the leaders. With the former, the picture is quite consistent. The latter, however, demonstrates that for retweets, the group of official accounts ranks higher—second rather than fourth—as compared to likes, while parody accounts rank lower—fourth. A cursory evaluation of this shift, based on content analysis, revealed an unusually high activity of automated Twitter accounts (i.e. bots), indicating an evidence of selective strategy of boosting certain topics.

The types of leaders also differ from one another in terms of their capacity to influence the content and sentiment of online conversations. To perform this task, the most critical metrics of individual tweets—likes and retweets—were queried from a sample of 3200 most recent tweets, authored by the leaders. These metrics were aggregated and the average number of “likes” and retweets per leader was calculated. Used independently, it provides a good estimate of the “average power of a tweet” of the given user, although it does not consider the issue of outliers—accounts with relatively short lifespan and

Table 30.3 Leaders’ impact metrics using all data collected or focused on the top 10k popular tweets

<i>Account type</i>	<i>All tweets</i>		<i>Sample of top 10K popular tweets</i>	
	<i>Av. Num. retweets</i>	<i>Av. Num. likes</i>	<i>Av. Num. retweets</i>	<i>Av. Num. likes</i>
Personal (302)	(2) 26.80	(2) 27.67	(3) 691.88	(2) 959.91
Media (86)	(3) 15.31	(3) 12.10	(1) 848.53	(3) 855.88
Official/Gov-t (49)	(4) 12.37	(4) 10.60	(2) 697.61	(4) 691.25
Parody (32)	(1) 98.35	(1) 150.74	(4) 664.49	(1) 1303.07

yet, quite high performance metrics. To address this, the maturity of accounts was estimated by multiplying the “tweet power” metric by the average number of tweets per day (1). Given the fact that all accounts in the leaders sample are real and used actively, the issue of automated content generation did not affect the overall calculations. Assuming that bots are less likely to be followed by leaders, the sample showed no evidence of the presence of unusually and/or suspiciously active accounts. Therefore, the average leader account generates approximately 16.03 (+/-2.3) tweets per day and the most active account, quite expectedly belonging to the media group, generates on average 170 tweets per day.

$$(1) \text{Average tweet power of Leader}_i = \left(\overline{\text{Favorites}_i} + \overline{\text{Retweets}_i} \right) * \overline{\text{NumTweetsPerDay}_i}$$

The overall list of candidate impact obtained by (1) was sorted from most to least impactful, and the overall list of 469 was broken down into quantiles. Figure 30.4 represents the breakdown of account types per quantile. Obviously, the parody group accounts generate content ordinary citizens are eager to react to: 53.1% of such accounts in the sample appeared in the first quantile. Approximately a quarter of personal accounts demonstrate the tendency to generate highly resonant content (23.5% in the first quantile). Interestingly, this most populous group is almost evenly distributed. Official accounts follow a somewhat normal distribution, peaking in the third quantile, hence generating relatively impactful content. The discrepancy between this distribution and the high performance of these account types in the top 10,000 sub-sample (Table 30.2) raises the importance of future in-depth content analysis of messages produced by this group. On the one hand, this content could be artificially “boosted”; on the other, top “tweets” could actually discuss politically crucial issues and be genuinely shared alongside the network, which to some

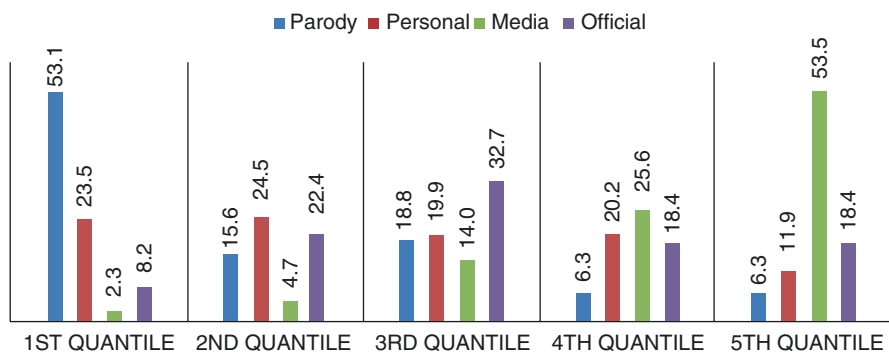


Fig. 30.4 Influence of leaders' content, distribution across quantiles

extent, supports the thesis of the bursty nature of Twitter networks (Myers and Leskovec 2014). Finally, and rather surprisingly, the most prolific group of media accounts tends to be distributed towards the lower part of the scale.

30.3.3.2 *Developing an Index of “Influence”*

The average tweet power metric for Twitter data, while indicative of certain patterns in ongoing and historical discussions, does not take into consideration network “influence” and the leaders’ ability to disseminate content throughout the network and in their particular communities. While the topic has gained significant attention in the research community, no established and widely accepted method of identifying Twitter influencers exists. Time-invariant approaches tend to compute influence on the basis of either centrality (network-dominated approach), or content impact (retweet-dominated approach). At the same time a combination of both methods could be quite productive. We propose a method utilizing network centrality and demonstrated ability to disseminate content.

Influence can be defined as the ability to seed discussions and spread content throughout the network. It can be seen as a derivative of two major parameters: the importance of content and its ability to meet the aspirations of ordinary users and the capacity of this content to spread through the network and be visible to a wide audience. The former is marked by users’ reaction to content, similar to the approaches taken in evaluating “influencers” above. The latter, on the other hand, evaluates the placement of the leader within a network or community, as a central placement creates a better opportunity to disseminate content amongst a wider audience. As such, we determine PageRank centrality on the “follow” relationship of Twitter, which is seen as both as an indicator of information-gathering, as well as social connection between two users, especially if it is reciprocated (Myers et al. 2014; Frederick et al. 2012).

Centrality is the most commonly used approach to determine the importance of nodes in a network (Livne et al. 2011; Romero et al. 2011). PageRank Centrality (Page et al. 1999), most famously used in Google search, assigns a probability distribution to the network, representing the chance of randomly picking a specific node. When applied to social networks, it allows the ranking of users by importance relative to each other. Centrality was combined with the aggregated average number of “likes” and “retweets” obtained from the 3200 tweets authored by each individual “leader.” Combining both parameters yields an index of identified leaders’ influence, which represents the potential to have an impact, rather than a *bona fide* substantiation of influence. Adopting the average tweet power metric (1), we multiply it by PageRank centrality of each candidate to generate “influence” index (λ).

$$(2) \text{Leaders' influence}_i = \text{Average tweet power of Leader}_i * \text{Centrality}_i$$

Introducing centrality and combining the data with leaders’ assigned *InfoMap* communities alters the observed distribution considerably. Firstly, the first two quantiles of the most influential Twitter users are comprised

predominantly of opposition accounts (Fig. 30.7, Annex B). In the first quantile, two-thirds (or 60%) of accounts belong to the opposition and only one-third to the pro-government community. A similar situation is observed in the second quantile, where 71% of accounts can be referred to as belonging to the opposition, and only 29% to a pro-government group. The first quantile included such popular opposition leaders as Aleksei Navalny, Leonid Volkov, Oleg Kashin, media outlets TV Rain (*Dožd'*), *Èho, Moskvy*, and *Meduza*, as well as highly influential parody accounts. The pro-government group, although outnumbered by its opponents, is represented by its most outspoken pundits (Vladimir Solovyev and Alexei Pushkov) and notable media sources (*RIA Novosti*, *Vesti News*). Interestingly, the most followed political accounts of Vladimir Putin and Dmitry Medvedev, although appearing in the top quantile, are located in the middle and in the end of it respectively. Secondly, the distribution of account types across quantiles is more flat, with a decline in proportion of parody and personal accounts in the first quantile and an increase in media and official accounts, and a gain in parody and media accounts at the expense of personal and official accounts, in the second quantile (Fig. 30.8, Annex B). This can be understood as indication that media and official accounts, while not impactful in terms of content, are central to the network and hence have a higher ability to distribute their content. The shift of a certain proportion of parody accounts from the first to the second quantile, as well as the relative decline of personal accounts, is a further validation of the presence of the “echo chamber.” As parody and personal content is usually popular in specific audiences, these accounts are not highly followed by opposing communities and hence do not share central position in the whole network.

Furthermore, such dominance of opposition accounts in the top half of the influence index speaks of the higher importance of this form of communication for the opposition and also supports evidence of the greater structuration and network sophistication from the network analysis. The opposition not only focuses on social media as its main form of reaching the audience but also emphasizes the role of opinion leaders. In this regard, Alexei Navalny is the major actor and the greatest influencer not only within his own political community, but also in the entire network. Pro-government pundits, like Vladimir Solovyev and Alexei Pushkov outperform their own formal leaders in terms of influence in the virtual community, and accounts of traditional federal mass media are instrumental in the dissemination of the pro-government content. This establishes a new framework of evaluating Russian political Twitter, which is quite different from Kelly et al. (2012) in terms of network structure and from Greene (2018) in terms of content.

30.3.3.3 Cross-Validation of the Proposed “Influence” Index

The proposed index (2) requires further testing and validation. Given the nature of the research topic, where outcomes are easily predictable on the basis of traditional theories and concepts of Russian politics, the best way to test reliability of a new instrument would be the utilization of another approach. Given the fact that this new method is a derivative of major other influence indicators,

reusing them would result in unfavorable procedural overlap and, thus, similarity of outcomes. To overcome this issue, and avoid complex dynamic methods, this research adapts the principle that utilized the Hirsch index (h-index) of academic impact (3).

Hirsch is rather unexpectedly suitable and productive for measuring leaders' performance on Twitter, and even overcomes deficiencies visible in the context of scholarly work. Firstly, leaders on Twitter are akin to scholars in academia, producing content aimed at specific audiences and seek endorsement for their work in terms of citations or "likes" and "retweets." Secondly, both academic papers and blog messages increase their value through references, with the growth being well documented and easily accessible. Thirdly, academics and leading bloggers both tend to increase their visibility by producing the maximum possible high-quality content. Moreover, the ample quantity of blog posts overcomes the limitations of academic work, where the number of contributions is usually lower.

$$(3) \text{Leaders' influence with Hirsch}_i = \text{hirsch method}(\text{Favorites} + \text{Retweets}) \\ * \text{Centrality}_i$$

Therefore, the use of the h-index seems justified, as it addresses the issue of outliers (i.e. highly popular tweets) as well as the lifespan of accounts (immature, yet highly popular accounts) and provides a weighted rank of significant contributions. To put it simply, the h-index algorithm finds an "ideal point" between the number of contributions and their relative popularity, which for Twitter can be considered as the sum of "likes" and "retweets" for each users' post (*hirsch method(Favorites + Retweets)*). All leaders were ranked according to the obtained indices and the resultant list was compared with the ranked list of leaders, obtained through the index method proposed by this research. Spearman's rank correlation coefficient (ρ) was utilized to establish whether both methods were concordant. It demonstrated a high correlation coefficient of 0.69 between the proposed influence index (2) and the modified h-index (3). Notably, this coefficient was calculated when the h-index did not refer to the centrality parameter of each leader account. Including the centrality indicator increased the correlation coefficient to 0.80.

As a ranking algorithm, the h-index provides a useful method for establishing the most influential contributors and can be used for ranking leaders. It also confirms the validity of the proposed time-invariant influence rank (2). As any other methods, the h-index for Twitter is not without deficiencies and potentially may not be used for samples where leaders are highly popular and produce a large quantity of tweets. As the Twitter REST API limits access to 3200 most recent posts, the h-index will not be able to produce an index higher than the quantity of posts. Yet in the case of current measurements of Russian Twitter, this issue was not a problem, as the most popular user—Alexei Navalny—scored only 902 points on the scale. Furthermore, both indices (2

and 3) are consistent and consonant with common wisdom that the actual disposition of actors and organizations within the political arena should be correlated with their political influence.

30.4 BEYOND THE SCORE: CROSS VALIDATION OF DETECTED PATTERNS

30.4.1 Further Validating “Echo Chambers”

Focusing on intra- and cross-community conversation, we observed homophilous conversation patterns between the various communities as users tended to share content with like-minded individuals inside their own community, most especially between the pro-government and Opposition groups. Nominal homophily however can be misleading as users in small communities are more likely to converse across community lines simply because their community is small, and users in large communities are unlikely to converse outside their group due to its relative proportion. We adopt a method developed by Currarini et al. (2007) to validate the nominal homophily observed. Specifically, nominal homophily, or the proportion of conversations a community has within itself (H_i) is compared to its relative size within the network (w_i).

$$(4) H_i = w_i; (5) H_i > w_i; (6) H_i < w_i$$

Baseline homophily (4) occurs when the proportion of user conversations within a community equals the relative size of the community, indicating that on aggregate, users in that community show no special preference or bias for their own friends. Inbreeding homophily (5) indicates that users are biased and converse more often within their own group than is expected on the basis of its relative demographic size. Finally, if a community shows heterophilous patterns (6), the number of conversations within the group will be less than the relative size of the group.⁹ To enable comparisons between communities of various sizes and different conversation types on Twitter, we standardize homophily indicators for (7) baseline homophily, (8) inbreeding homophily, and (9) heterophily.

$$(7) \frac{H_i}{w_i} = 1; (8) \frac{H_i}{w_i} > 1; (9) \frac{H_i}{w_i} < 1$$

Investigating standardized homophily indicators ($\frac{H_i}{w_i}$), we find that each community demonstrated strong inbreeding homophily (Table 30.4). Interestingly, the non-systemic opposition is more homophilous than the pro-government community. A few communities, such as communities 3 and 4, while quite small, demonstrated excessively high standardized homophily indicators.

Table 30.4 Relative community sizes and standardized homophily indicators

	<i>Com 0</i> [<i>pro-government</i>]	<i>Com 1</i>	<i>Com 2</i> [<i>opposition</i>]	<i>Com 3</i>	<i>Com 4</i>
Relative size	22.7%	9.3%	6.2%	0.8%	0.4%
$\frac{H_i}{w_i}$ (retweets)	3.96	8.07	12.53	95.66	175.80
$\frac{H_i}{w_i}$ (mentions)	3.56	8.81	10.74	111.16	0.00

30.4.2 *Makeup of Two Main Political Communities and Their Reactions to Political Events*

Individually assessing the pro-government community, we observe that it was by far the largest political group, always actively participating in all events. The community displayed strong pro-Putin, pro-government, anti-western (including anti-US [United States], anti-Europe and anti-Ukraine), and, in some instances even nationalist sentiment. While sometimes a little critical of the regime, the users in this community (community 0 or purple in Fig. 30.1 above) generally disseminated information in line with a patriotic narrative and demonstrated two patterns of Twitter use. If the informational event was not negative to Russia or the government, such as the two-year anniversary of the accession of Crimea as part of the Russian Federation, then reactions were usually event-specific and generally positive. However, if the informational event was inherently negative to the government, reaction was usually split between a certain proportion of anti-government content, and neutral or positive pro-government reaction. In certain cases, a pattern is evident whereby positive content was coordinated around specific keywords that were trending negatively in order to coopt the term and spin it positively, distracting the conversation to unrelated pro-government content.

Reactions to Prime Minister Medvedev's comment of "there is no money, but you hang in there" to Crimean pensioners, posted to YouTube on May 23, 2016, demonstrate these two patterns well (Fig. 30.5). While some users derided the Prime Minister's comments, factual and neutral reactions were quite prevalent. A large amount of disseminated content focused on unrelated positive topics to distract and mitigate the initial negative reaction inside the community. Two stories were widely mentioned on May 23 and 24 focusing on specific keywords. The first focused on the word "money" by distributing a story on the Prime Minister promising to find money for museums in Crimea. The second focused on the terms "economy" and "investment," disseminating content about the release of a government plan, approved by the Prime Minister, aiming to increase domestic demand for the products of Russian chemical and petrochemical industries. Other reactions also included the factual reporting of the Prime Minister's comments or presenting the information in a neutral fashion, with tweets such as "Medvedev admitted that there is no money to index pensions."¹⁰ Interestingly, such neutral posts usually did not include links and were composed of just text.

The solid line indicates the number of tweets on an hourly basis (right axis) in the pro-government community, and the dashed line indicates the

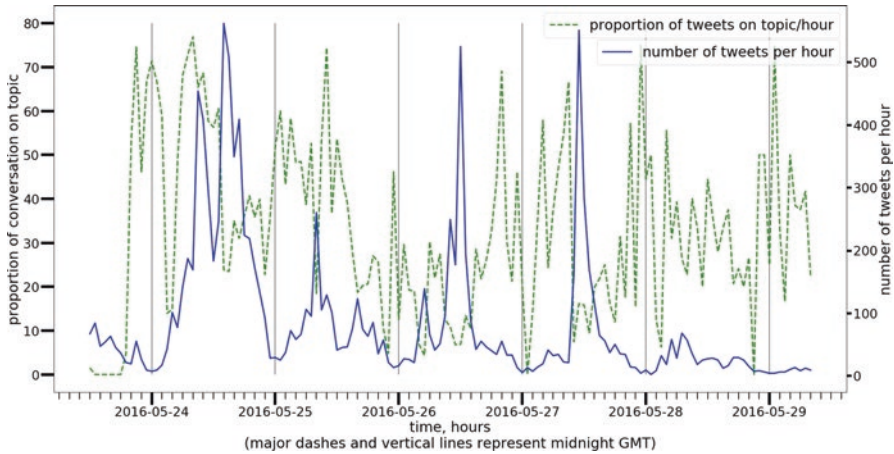


Fig. 30.5 Pro-government community reaction to Medvedev’s comment to pensioners in Crimea

proportion of conversation (i.e. of the tweets and retweets made during that hour) that had to do with Medvedev’s comment (left axis). Tweets containing “money,” “hang in there,” “pensioners,” “have a good day” (“*deneg*,” “*den’gi*,” “*deržites’*,” “*pensii*,” “*pensij*,” “*nastroenie*”) were used to calculate the proportion. Specifically, lemmatized words in each tweet were checked against the lemmas of desired keywords. Tweets that contained keywords known to be used in the counter strategy were excluded.

The second of the two main political groups, the Opposition community (community 2, marked teal in Fig. 30.1 above), was on average 3.5 times smaller. Its users displayed negative, ironic, and critical assessments of the Russian government and also disseminated Ukrainian-friendly, pro-US, and pro-Western content. While users in this community also shared content on liberal values, such as opposition to authoritarian government or support for democracy, a majority focused on vilifying the government, with users spreading negative memes or ridiculing government strategies or statements. Indeed, particularly virile ridicule and even contempt of the government followed Medvedev’s comments in Crimea. The community is also made up of a sizable proportion of Russian-speaking Ukrainians, which seemed to influence how the group reacts to informational events. Indeed the Savchenko affair and the Eurovision contest, both highly interrelated with Ukrainian politics, are the largest samples of captured Opposition group users, with the former being the largest by number of tweets and the latter, the largest in quantity of engaged users.

Similar to the pro-government community, content shared within the Opposition group followed a dual pattern. If the informational event was neutral or negative to the government, and hence in line with community expectation, users either discussed the topic in a neutral fashion or spread content negative to the government. However, when the informational event ran counter to community expectations, then reaction was split as users reacted in

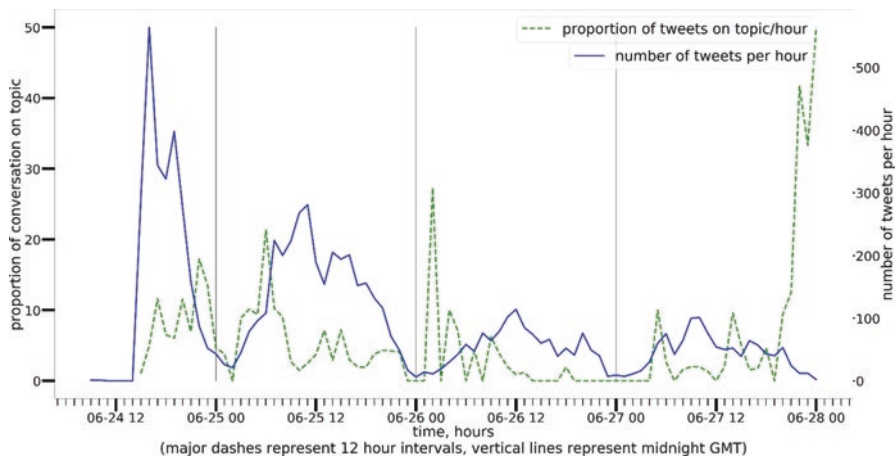


Fig. 30.6 Opposition community reaction of disbelief to Belykh's guilt

different ways. The three-pronged reaction following the arrest of the liberal governor of the Kirov Oblast, Nikita Belykh, well demonstrates this trend. Factual and neutral tweeting was predominant; however, genuine shock and disbelief, often including statements that Belykh was set up, was also prevalent (see Fig. 30.6). Finally, a third opinion expressed by the community was that of anti-Belykh statements, believing that he betrayed the liberal movement by becoming a systemic politician.

Represents data from June 25 to June 28. The solid line indicates the number of tweets per hour (right axis) and the dashed line indicates the proportion of conversation believing Belykh was set up (left axis). Tweets containing lemmatized words including “setup,” “don’t believe” and “provocation” (“*podstava*,” “*podstavili*,” “*podstavit*,” “*podstav*,” “*ne verú*,” “*poverit*,” “*provokaciá*,” “*provocirovali*”) were used to calculate the proportion.

Comparing the standardized rate of tweets per hour between the two main communities, we see that the pro-government group reacted very differently than the Opposition group to several events, most notably during the Medvedev event (see Medvedev Chart in Annex C).¹¹ The initial spike of tweeting activity in mid-day on May 24 lagged by a few hours the initial and larger reaction by the Opposition group, indicating that the pro-government group was less likely to immediately react to the negative information. The secondary and larger spike of activity in the latter half of the day is particularly interesting, given its size and the content shared during it was mostly not about Medvedev’s comment at all. Comparing the (nominal) number of tweets every two hours with the proportion of the tweets that have to do with Medvedev’s comments, we see that the conversation shifted to discussing other topics during this second spike (see Fig. 30.5). Following this, the proportion of unrelated content on Medvedev continued to dominate discussions within the community, with the proportion mentioning the pensioner comment mostly remaining in the

20–30% range. Moreover, given that the keyword used to collect this sample was “Medvedev,” we hypothesize that this pattern of distraction specifically had to do with positively portraying the Prime Minister.

30.4.3 *Finding Bots Within Russian Twitter*

SNA is undeservedly neglected in the context of the mainstream topic of “bot” and “troll” impact on Russian “online” political discussions. Recent research points that the proportion of content created on Russian Twitter has a 50% probability of being produced by “bots” (Stukal et al. 2017). As stated elsewhere (Murthy et al. 2016), bots can have an impact on simple indicators such as follower counts or hashtag boosting. This may impact users who follow others indiscriminantly and network metrics as a whole, which assess all tweets in the network without taking into account the structure of social networks (Ferrara 2018). As ordinary users tend not to follow “bots” and segregate themselves into isolated “echo-chambers,” bots are likely to be segregated to isolated communities that have little influence on real politically engaged users.

To evaluate the potential impact and ensure the validity of our findings, we apply a three-method strategy to measure the prevalence of bots within the identified network structure. First, we evaluate the proportion of duplicate and highly similar content created in each community, as bots are known to repeat (not retweet) identical tweets (Lawrence 2015). As tweets can have very minor purposeful variation, applied to them by bot designers, such as adding an extra hashtag or changing the beginning of the tweet, we compare the similarity of tweets by excluding tweet extremities. Secondly, to validate the results of the first method, we qualitatively assess a sub-sample of tweets shared in suspect communities. Finally, we apply the popular *Bot or Not* method, also known as botometer, to score the likelihood of each user in our samples to be a “bot” (Davis et al. 2016), a feature-based method that evaluates a set of behaviors of a Twitter account and assigns it a score (probability) of being a “bot” (Ferrara 2018). A tried and relatively accurate approach (Yang et al. 2019), it is appropriate for cross-validating other methods utilized.

We find that in the two main political communities, the proportion of duplicate and similar tweets varied by event; however, the pro-government community demonstrated a much larger proportion of both in all events (Table 30.5). For instance, in the Medvedev sample, 43% of tweets are very similar in the pro-government community, many spreading the positive new story of government support of Russia’s petrochemical industry on May 24. The low proportion of duplicate content in the opposition community is intriguing, as is the relatively high proportion of similar content. A possible explanation could be that the opposition is a more dynamic group of users who follow more sophisticated bots that utilize more complex natural language or image methods to spread content. We propose this fascinating puzzle as a question for further research in the area.

Outside the main political communities, qualitative and duplicate/similarity analysis revealed that many visually segregated communities were made up to a very large degree of “bots,” or at least accounts sharing very similar content

Table 30.5 Duplication and similarity of content in two main communities (% identical; % similar)

<i>Event</i>	<i>Pro-government community</i>	<i>Opposition community</i>
Belyh	12.4%; 28.2%	0.1%; 6.6%
Crimea	1.2%; 5.1%	NA; 2.2%
Eurovision	10.8%; 27.9%	0.3%; 13.3%
Medvedev	11.9%; 43.1%	1.1%; 19.5%
Savčenko	11%; 31.5%	0.2%; 24.5%
Turkey	12.7%; 33%	0.1%; 16.4%

Table 30.6 Sample of “bot” communities detected (% identical; % similar)

<i>Event</i>	<i>Community 3 (light green)</i>	<i>Community 7 (dark orange)</i>	<i>Community 8 (dark purple)</i>	<i>Community 21 (yellow)</i>	<i>Community 43 (red)</i>
Belyh	53%; 54%	24%; 24%	NA; NA	60%; 60%	NA; 27%
Crimea	NA	NA	NA	NA; 96%	NA; NA
Eurovision	43%; 46%	83%; 83%	3%; 16%	73%; 74%	NA; 11%
Medvedev	31%; 36%	61%; 61%	3%; 19%	44%; 47%	NA; 13%
Savčenko	27%; 32%	64%; 64%	3%; 12%	71%; 72%	2%; 12%
Turkey	46%; 49%	24%; 24%	NA; NA	57%; 57%	8%; 36%

(Table 30.6). The more unsophisticated groups posted identical or very similar content, including as high as 83% of all tweets for an event (community 7, dark orange). Others showed more complex approaches, such as tweeting news headlines or factual statements, with or without a corresponding link in the tweet. Interestingly, when links were present, they often pointed to Yandex News or even more commonly to heterodox news or blogging sites. While the information captured in the samples of the six events for these communities were often political, the public timelines of the “bot” accounts often included completely unrelated content, such as for commercial purposes and advertising. We assume that these communities of “bots” are owned by marketing or consulting firms and tweet specific content based on the requirements of their clients, without any particular impact on actual political discussions.

Evaluating the average automation probability of users in each community, as reported by botometer, reinforces our findings (Table 30.7). Communities of real users had on average low probability of being automated, with a relatively small proportion of users removed or suspended. By the same token, communities, previously identified as highly likely as being “botnets” or having a large prevalence of “bots” (Communities 3, 7, 8 and others) had much higher average probability of being automated. Moreover, a large proportion of accounts in these communities have since been suspended by Twitter. Recently the company expanded its activities to diminish the impact of bots and trolls by suspending multiple accounts.¹² The results of these actions reinforced our findings, as the proportion of suspended accounts in the communities we identified as botnets was much higher than in real user communities. Indeed, entire

Table 30.7 Botometer (Bot or Not) results (average universal probability of automation; proportion of accounts no longer present two years after samples were originally collected)

<i>Event</i>	<i>Pro-government (community 0)</i>	<i>Opposition (community 2)</i>	<i>Community 3</i>	<i>Community 7</i>	<i>Community 8</i>
Belykh	11.2%; 15%	5.8%; 10.1%	49.8%; 37.1%	50.4%; 4.8%	NA; 100%
Crimea	7.8%; 12.4%	7.6%; 9.7%	54.7%; 50%	NA	NA
Medvedev	15.8%; 17%	7.8%; 10.5%	53.6%; 38.1%	NA	37.8%; 97%
Savchenko	11.7%; 15.1%	8.8%; 9.8%	57.2%; 32.6%	45.6%; 3.6%	28.6%; 98.2%

“botnets” have all but been suspended by Twitter in the two years since the samples were originally collected (see Table 30.7). On the other hand, some remain relatively untouched. From all above indicators, we conclude that “bots,” while undeniably highly numerous and often verbose on Russian Twitter, are often segregated to isolated mini communities that have little impact on real politically engaged users. Real political communities do likely possess a certain proportion of bots, however, as identified in the literature (Kollanyi 2016; Murthy et al. 2016), these bots are likely to be complex and highly sophisticated, making their study challenging but their potential impact on shifting real conversations greater.

30.5 CONCLUSION

Data on public and political engagement of Russian citizens, active political discussions and debates, and even protest coordination activities, are readily available to researchers studying Russian politics. This chapter illustrates how SNA can be instrumental and ineluctable in evaluating key research hypotheses utilizing such data. Using six resonant political discussions collected from Twitter over the summer of 2016, we validate multiple political and sociological theories important for Russian studies. Firstly, we find that Russian online society is divided among the same ideological lines as the public sphere, representing two distinct and consistent communities of users, one supporting the Kremlin, and a “non-systemic” opposition that opposes it. Secondly, we validate the presence of “echo-chambers” in Russian social networks, identifying polarization between the two main political groups. Thirdly, we observe that “influencers” on Russian Twitter are not generally traditional political elites, but “interesting” and highly informative users such as that of famous pundits, parody accounts, or news sources. Furthermore, given regular users’ isolation into self-created separate ideological communities as well as further mini-communities, we comment that the expected impact of information control strategies by the government are likely quite limited.

We obtain our results through the application of a thorough and holistic approach in evaluating the network structure, focusing on three levels of analysis. Macroscopic methods, such as community detection and network visualization, supports the evaluation of the “overall” picture of online Russian political

society. Mesoscopic methods validate the detected structures and provide insight into the sub-structure of each detected community. Finally, microscopic methods identify the “influential” users who are able to widely disseminate content and impact political discussions.

We find that SNA is also an economical method to detect “bots” in social networks and evaluate their impact on real political users. Currently a hotly debated topic both internationally (Murthy et al. 2016; McKelvey and Dubois 2017; Ferrara 2018) and in Russia (Stukal et al. 2017; Lawrence 2015), our findings demonstrate that the impact of “bots” on Russian social networks is likely quite negligible. We find that while numerous and often verbose, “bots” are mostly isolated to mini-communities far removed from real politically engaged users, and as such are unlikely to impact real political discussions. We conclude by noting that the efficiency of SNA in extracting real and valid structures in Russian social networks makes it a prerequisite and fundament for the application of further advanced methods, such as topic or sentiment analysis, when studying Russian politics (for more on sentiment analysis, see Chap. 28).

ANNEX A: POLARIZATION OF COMMUNITIES

Polarization of communities is taken from an aggregate of all six events. Note that as over 1000 groups were detected, only the most numerically relevant are shown. Percentages represent proportion of the total for each community (the “All” category).

Table 30.8 Polarization of communities: retweets

<i>Retweets</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>8</i>	<i>All</i>
0	89.7%	0.4%	3.0%	0.0%	0.0%	0.0%	0.0%	194,393
1	2.7%	75.4%	0.3%	0.0%	0.0%	0.0%	0.0%	33,392
2	6.8%	0.0%	77.9%	0.0%	0.0%	0.0%	0.0%	54,672
3	15.3%	0.4%	0.5%	75.3%	0.0%	0.0%	0.0%	837
4	11.9%	1.4%	0.7%	1.4%	75.9%	0.0%	0.0%	286
5	50.5%	1.0%	7.5%	0.5%	0.0%	13.0%	0.0%	200
8	1.1%	0.0%	0.0%	0.0%	0.0%	0.0%	19.5%	87
All	196,233	33,698	53,563	793	232	138	22	350,616

Table 30.9 Polarization of communities: mentions

<i>Mentions</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>5</i>	<i>All</i>
0	80.7%	0.3%	8.8%	0.0%	0.0%	12,114
1	2.9%	82.3%	0.5%	0.0%	0.0%	1405
2	21.0%	0.3%	66.8%	0.0%	0.0%	3312
3	7.5%	0.0%	1.3%	87.5%	0.0%	80
5	35.0%	0.0%	15.0%	0.0%	0.0%	20
All	11,584	1456	3671	70	1	21,691

Table 30.10 Polarization of communities: replies

<i>Replies</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>All</i>
0	85.3%	0.1%	6.5%	5354
1	2.0%	83.7%	0.0%	245
2	24.3%	0.1%	63.7%	1172
All	5423	250	1206	8083

ANNEX B: MICROSCOPIC

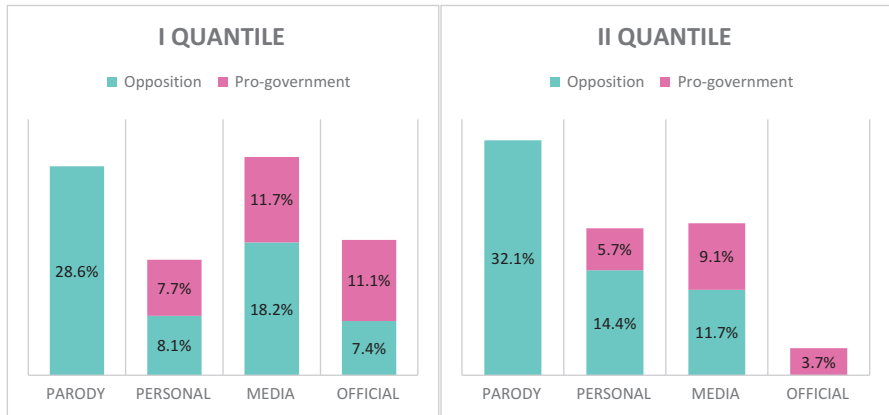


Fig. 30.7 First and second quantile breakdown of account types by political communities

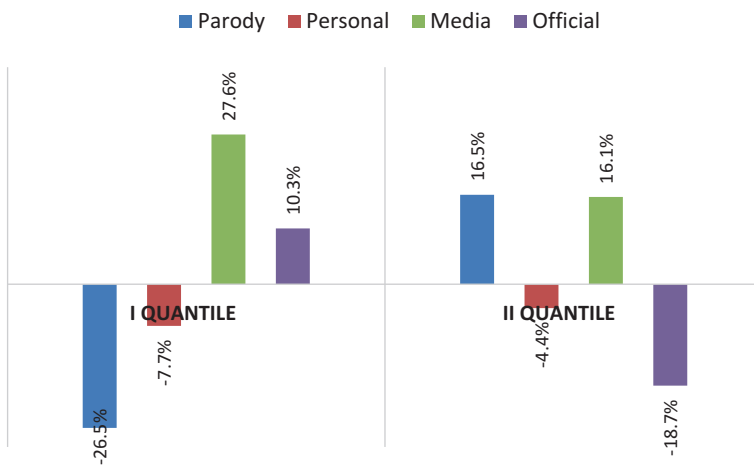


Fig. 30.8 Change in influence of leaders' accounts when centrality complements the impact of their distributed content

ANNEX C: TWEET PATTERNS FOR THE TWO MAIN COMMUNITIES, TWEETS PER HOUR BY EVENT

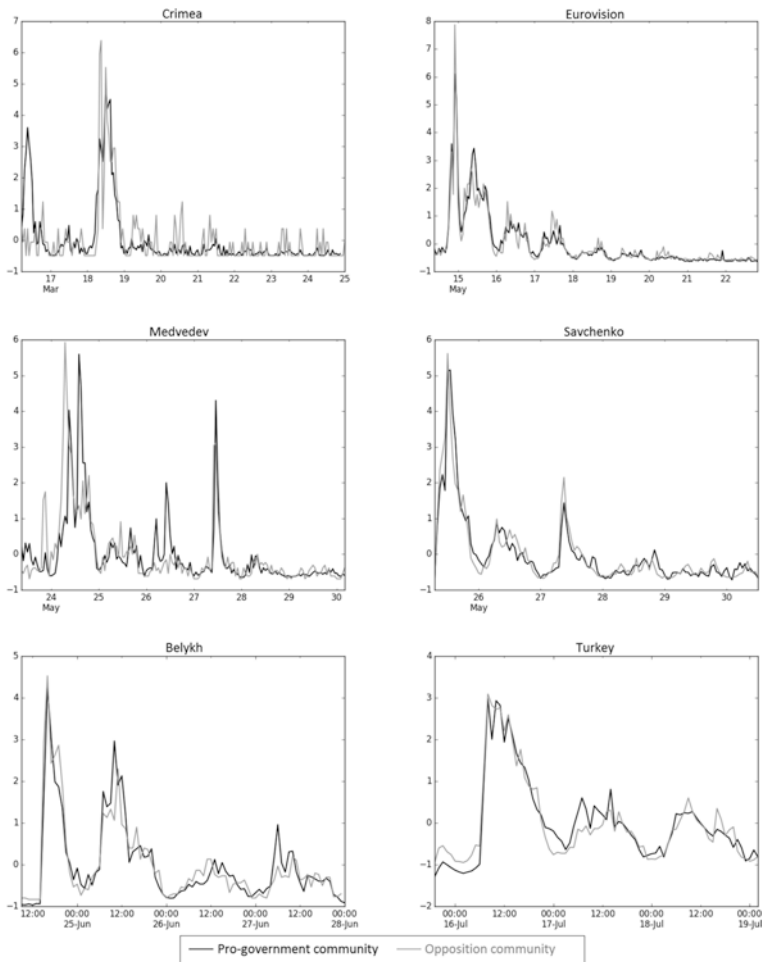


Fig. 30.9 Tweet patterns for the two main communities, tweets per hour by event

NOTES

1. Compared to 46.6M users a month for VK and an internet penetration of 109.5M. Source: *Translate Media* (<https://www.translatemedia.com/translation-services/social-media/russian-social-media/>).
2. According to the *Statista.com*. Source: <https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/>.
3. In 2016, amongst the top 100, 17 accounts belonged to politicians, government organizations or media personalities who comment mostly on political

- issues, and a further 17 were of news media. In 2017, the pattern was similar with 15 and 16 respectively. Prime Minister Dmitry Medvedev has remained the most followed politician and always either tops or is at the top of the list. Although the trend has shifted towards popularization of non-political accounts, as of 2019, the nominal politicization of Twitter still remains quite high with 28 politically relevant accounts among the top 100. In comparison, German, French, UK, and American politicians' positions are less representative in the top 100, as these segments are largely dominated by the non-political (celebrities, media and sport personalities) opinion leaders.
4. Lancichinetti and Fortunato (2009) provides a good overview of available methods for various data and network constraints.
 5. As a test case, we amalgamated the six samples into one master network, which yielded a modularity of 0.551—a relatively strong statistic of network stratification.
 6. For introduction and application of methods in Network Analysis, see Zinoviev (2018).
 7. See Twitter's Developer documentation for available streaming parameters and their limitations: <https://developer.twitter.com/en/docs/tweets/filter-real-time/guides/basic-stream-parameters.html>.
 8. For instance, the #followback hashtag is commonly used by users to increase their number of followers.
 9. See Currarini et al. (2007) for in depth explanation.
 10. Note that this type of tweet was repeated by many users with slight variation in language.
 11. Given the difference in the sizes of the pro-government and Opposition communities, a comparison of their participation in events was performed by transforming the nominal tweets per hour scale into a standardized one using corresponding t-scores.
 12. In July 2018, the Washington Post reported that Twitter suspended 70M accounts in May and June. See Timberg and Dwoskin (2018) for more information.

REFERENCES

- Agarwal, S.D., W.L. Bennett, C.N. Johnson, and S. Walker. 2014. A Model of Crowd Enabled Organization: Theory and Methods for Understanding the Role of Twitter in the Occupy Protests. *International Journal of Communication* 8: 646–672.
- Bakshy, Eytan, Jake M. Hofman, Winter Mason, and Duncan J. Watts. 2011. Everyone's an Influencer: Quantifying Influence on Twitter. *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining SE – WSDM '11*: 65–74. <https://doi.org/10.1145/1935826.1935845>.
- Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. 2015. Exposure to Ideologically Diverse News and Opinion on Facebook. *Science* 348 (6239): 1130–1132. <https://doi.org/10.1126/science.aaal160>.
- Barash, Vladimir, and John Kelly. 2012. Saliency vs. Commitment. Dynamics of Political Hashtags in Russian Twitter. *The Berkman Center Research Publication* 9: 0–16. <https://ssrn.com/abstract=2034506>.
- Barberá, Pablo, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau. 2015. Tweeting from Left to Right: Is Online Political Communication More Than

- an Echo Chamber? *Psychological Science* 26 (10): 1531–1542. <https://doi.org/10.1177/0956797615594620>.
- Bastian, Mathieu, Sebastien Heymann, and Mathieu Jacomy. 2009. Gephi: An Open Source Software for Exploring and Manipulating Networks. In *Third International AAAI Conference on Weblogs and Social Media*, 361–362. <https://doi.org/10.1136/qshc.2004.010033>.
- Borgatti, Stephen P., Ajay Mehra, Daniel J. Brass, and Giuseppe Labianca. 2009. Network Analysis in the Social Sciences. *Science* 323 (5916): 892–895. <https://doi.org/10.1126/science.1165821>.
- Brandes, Ulrik, Daniel Delling, Marco Gaertler, Robert Görke, Martin Hoefer, Zoran Nikoloski, and Dorothea Wagner. 2007. On Finding Graph Clusterings with Maximum Modularity. In *Graph-Theoretic Concepts in Computer Science*, 121–132.
- Breuer, Anita, Todd Landman, and Dorothea Farquhar. 2015. Social Media and Protest Mobilization: Evidence from the Tunisian Revolution. *Democratization* 22 (4): 764–792. <https://doi.org/10.1080/13510347.2014.885505>.
- Coleman, James. 1990. *Foundations of Social Theory*. Cambridge, MA: Belknap.
- Colleoni, Elanor, Alessandro Rozza, and Adam Arvidsson. 2014. Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data. *Journal of Communication* 64 (2005): 317–332. <https://doi.org/10.1111/jcom.12084>.
- Currarini, Sergio, Matthew O. Jackson, and Paolo Pin. 2007. An Economic Model of Friendship: Homophily, Minorities and Segregation. *Ssrn* 20. <https://doi.org/10.2139/ssrn.1021650>.
- Davis, Clayton A., Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. *BotOrNot: A System to Evaluate Social Bots*, 4–5. <https://doi.org/10.1145/2872518.2889302>.
- Earl, Jennifer, Heather McKee Hurwitz, Analia Mejia Mesinas, Margaret Tolan, and Ashley Arlotti. 2013. This Protest Will Be Tweeted: Twitter and Protest Policing During the Pittsburgh G20. *Information Communication and Society* 16 (4): 459–478. <https://doi.org/10.1080/1369118X.2013.777756>.
- Ferrara, Emilio. 2018. Measuring Social Spam and the Effect of Bots on Information Diffusion in Social Media. In *Complex Spreading Phenomena in Social Systems*, 229–255. Springer. https://doi.org/10.1007/978-3-319-77332-2_13.
- Frederick, Evan L., Choong Hoon Lim, Galen Clavio, and Patrick Walsh. 2012. Why We Follow: An Examination of Parasocial Interaction and Fan Motivations for Following Athlete Archetypes on Twitter. *International Journal of Sport Communication* 5 (4): 481–502.
- Fuchs, Christopher. 2014. *OccupyMedia!: The Occupy Movement and Social Media in Crisis Capitalism*. Alresford, UK: Zero Books.
- Gel'man, Vladimir. 2015. Political Opposition in Russia: A Troubled Transformation. *Europe-Asia Studies* 67 (2): 177–191. <https://doi.org/10.1080/09668136.2014.1001577>.
- Golos. 2011. *Pervoe Zaávlenie asociacii 'GOLOS' po rezul'tatam dolgosrochnogo nablúdeniá boda izbiratel'nyh kampanij mestnyh, regional'nyh i vyborov deputatov Gosudarstvennoj dumy Rossii, naznačennyh na 4 dekabrâ 2011 g. Étapy vydviženíá i registracii*.
- Greene, Samuel. 2013. Beyond Bolotnaia. *Problems of Post-Communism* 60 (2): 40–52. <https://doi.org/10.2753/PPC1075-8216600204>.

- . 2018. Twitter and the Russian Street: Memes, Networks and Mobilization. *Center for New Media and Society Working Paper 1*.
- Halberstam, Yosh, and Brian Knight. 2015. Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter. *Journal of Public Economics* 143: 73–88. <https://doi.org/10.1016/j.jpubeco.2016.08.011>.
- Jacomy, Mathieu, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. 2014. ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *PloS One* 9 (6): e98679.
- Jensen, Michael. 2018. Russian Trolls and Fake News: Information or Identity Logic? *Journal of International Affairs* 71 (1.5): 115–124.
- Jungherr, Andreas. 2014. The Logic of Political Coverage on Twitter: Temporal Dynamics and Content. *Journal of Communication* 64 (2): 239–259. <https://doi.org/10.1111/jcom.12087>.
- Kelly, John, Vladimir Barash, Karina Alexanyan, Bruce Etling, Robert Faris, Gasser Urs, and John G. Palfrey. 2012. Mapping Russian Twitter. *The Berkman Center Research Publication* 760 (2010): 1–60. <https://doi.org/10.5465/amr.2011.0193>.
- Kollanyi, Bence. 2016. Where Do Bots Come from? An Analysis of Bot Codes Shared on GitHub. *International Journal of Communication* 10 (June): 4932–4951.
- Koren, Yehuda. 2003. On Spectral Graph Drawing. In *International Computing and Combinatorics Conference*, 496–508. Berlin.
- Lancichinetti, Andrea, and Santo Fortunato. 2009. Community Detection Algorithms: A Comparative Analysis. *Physical Review E* 80 (9): 1–12.
- Lawrence, Alexander. 2015. Social Network Analysis Reveals Full Scale of Kremlin's Twitter Bot Campaign. *Globalvoices* 2015. <https://globalvoices.org/2015/04/02/analyzing-kremlin-twitter-bots/>.
- Livne, Avishay, Matthew P. Simmons, Eytan Adar, and Lada A. Adamic. 2011. The Party Is Over Here: Structure and Content in the 2010 Election. *Association for the Advancement of Artificial Intelligence* 161 (3): 201–208. <https://doi.org/10.1007/s00024-003-2459-0>.
- Lonkila, Markku. 2012. Russian Protest On- and Offline: The Role of Social Media in the Moscow Opposition Demonstrations in December 2011. *FIIA BRIEFING PAPER* 98.
- Lotan, Gilad, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce, and Danah Boyd. 2011. The Arab Spring | The Revolutions Were Tweeted: Information Flows During the 2011 Tunisian and Egyptian Revolutions. *International Journal of Communication* 5 (0): 31. <https://ijoc.org/index.php/ijoc/article/view/1246>.
- Maslinsky, Kirill, Sergey Koltcov, and Olessia Koltsova. 2013. Changes in the Topical Structure of Russian-Language LiveJournal: The Impact of Elections 2011. *Ssrn*. <https://doi.org/10.2139/ssrn.2209802>.
- McKelvey, Fenwick, and Elizabeth Dubois. 2017. Computational Propaganda in Canada: The Use of Political Bots. *The Computational Propaganda Project Working Paper Series*.
- Murthy, Dhiraj, Alison B. Powell, Ramine Tinati, Nick Anstead, and Susan J. Halford. 2016. Bots and Political Influence: A Sociotechnical Investigation of Social Network Capital. *International Journal of Communication* 10 (June): 4952–4971.

- Myers, Seth, and Jure Leskovec. 2014. The Bursty Dynamics of the Twitter Information Network. In *Proceedings of the 23th International Conference on World Wide Web (WWW)*, 913–924. <https://doi.org/10.1145/2566486.2568043>.
- Myers, Seth A., Aneesh Sharma, Pankaj Gupta, and Jimmy Lin. 2014. Information Network or Social Network? The Structure of the Twitter Follow Graph. *International World Wide Web Conference*: 493–498. <https://doi.org/10.1145/2567948.2576939>.
- Nagornyy, Oleg, and Olessia Koltsova. 2017. Mining Media Topics Perceived as Social Problems by Online Audiences: Use of a Data Mining Approach in Sociology. *Srvn*. <https://doi.org/10.2139/ssrn.2968359>.
- Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. “The PageRank Citation Ranking: Bringing Order to the Web.”
- Pallin, Carolina Vendil. 2017. Internet Control Through Ownership: The Case of Russia. *Post-Soviet Affairs* 33 (1): 16–33.
- Pfeffer, Jürgen, Katja Mayer, and Fred Morstatter. 2018. Tampering with Twitter’s Sample API. *EPJ Data Science* 7 (1): 50. <https://doi.org/10.1140/epjds/s13688-018-0178-0>.
- Reuter, Ora John, and David Szakonyi. 2013. Online Social Media and Political Awareness in Authoritarian Regimes. *British Journal of Political Science* 45 (1): 29–51. <https://doi.org/10.1017/S0007123413000203>.
- . 2015. Online Social Media and Political Awareness in Authoritarian Regimes. *British Journal of Political Science* 45 (1): 29–51.
- Roesen, Tine, and Vera Zvereva. 2014. Social Network Sites on the Runet. In *Digital Russia: The Language, Culture and Politics of New Media Communication*, ed. Ingunn Lunde, Michael Gorham, and Martin Paulsen, 72–87.
- Romero, Daniel M., Wojciech Galuba, Sitaram Asur, and Bernardo A. Huberman. 2011. Influence and Passivity in Social Media. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, 18–33.
- Rosvall, Martin, and Carl T. Bergstrom. 2008. Maps of Random Walks on Complex Networks Reveal Community Structure. *PNAS* 105 (4): 1118–1123. <https://doi.org/10.1103/PhysRevLett.100.118703>.
- Rosvall, Martin, D. Axelsson, and Carl T. Bergstrom. 2009. The Map Equation. *European Physical Journal: Special Topics* 178 (1): 13–23. <https://doi.org/10.1140/epjst/e2010-01179-1>.
- Sinclair, Betsy. 2016. Network Structure and Social Outcomes: Network Analysis for Social Science. In *Computational Social Science*, ed. R. Michael Alvarez, 121–139. Cambridge University Press.
- Spaiser, Viktoria, Thomas Chadefaux, Karsten Donnay, Fabian Russmann, and Dirk Helbing. 2017. Communication Power Struggles on Social Media: A Case Study of the 2011–12 Russian Protests. *Journal of Information Technology and Politics* 14 (2): 132–153. <https://doi.org/10.1080/19331681.2017.1308288>.
- Stukal, Denis, Sergey Sanovich, Richard Bonneau, and Joshua A. Tucker. 2017. Detecting Bots on Russian Political Twitter. *Big Data* 5 (4).
- Timberg, Craig, and Elizabeth Dwoskin. 2018. Twitter Is Sweeping out Fake Accounts Like Never Before, Putting User Growth at Risk. *The Washington Post* 2018.
- Towner, Terri L., and David A. Dulio. 2012. New Media and Political Marketing in the United States: 2012 and Beyond. *Journal of Political Marketing* 11 (1–2): 95–119.
- Tremayne, Mark. 2014. Anatomy of Protest in the Digital Era: A Network Analysis of Twitter and Occupy Wall Street. *Social Movement Studies* 13 (1): 110–126.

- Tucker, Joshua A., Jonathan Nagler, Megan Macduffee Metzger, Pablo Barberá, Richard Penfold-Brown, and Duncan Bonneau. 2016. Big Data, Social Media, and Protest: Foundations for a Research Agenda. In *Computational Social Science Discovery and Prediction*, ed. R. Michael Alvarez, 199–224.
- Tufekci, Zeynep, and Christopher Wilson. 2012. Social Media and the Decision to Participate in Political Protest: Observations from Tahrir Square. *Journal of Communication* 62 (2): 363–379.
- Waters, Richard D., and Jensen M. Williams. 2011. Squawking, Tweeting, Cooing, and Hooting: Analyzing the Communication Patterns of Government Agencies on Twitter. *Journal of Public Affairs* 11 (4): 353–363. <https://doi.org/10.1002/pa.385>.
- White, Stephen, and Ian McAllister. 2014. Did Russia (Nearly) Have a Facebook Revolution in 2011? Social Media’s Challenge to Authoritarianism. *Politics* 34 (1): 72–84.
- Wolfsfeld, G., E. Segev, and T. Sheafer. 2013. Social Media and the Arab Spring: Politics Comes First. *The International Journal of Press/Politics* 18 (2): 115–137. <https://doi.org/10.1177/1940161212471716>.
- Yang, Kai-Cheng, Onur Varol, Clayton A. Davis, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2019. Arming the Public with Artificial Intelligence to Counter Social Bots. *ArXiv Preprint ArXiv:1901.00912* 2019: 48–61. <https://doi.org/10.1002/hbe2.115>.
- Zherebtsov, Mikhail. 2019. Taking Stock of Russian EGovernment. *Europe-Asia Studies*: 1–29.
- Zinoviev, Dmitry. 2018. *Complex Network Analysis in Python: Recognize-Construct-Visualize-Analyze-Interpret*. Pragmatic Bookshelf.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

