DOI: http://doi.org/10.25728/cs.2021.3.2

INFORMATION COMMUNITIES IN SOCIAL NETWORKS. PART III: APPLIED ASPECTS OF DETECTION AND ANALYSIS[1]

L.M. Boiko[2], D.A. Gubanov[3], and I.V. Petrov[4]

Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia

[2]✉ boiko.lilia@gmail.com, [3]✉ dmitry.a.g@gmail.com, [4]✉ zyxzy@protonmail.ch

Abstract. This paper overviews the empirical studies of the formation and detection of information communities in social networks. In parts I and II of the survey, we outlined the concept of an information community and considered the relevant mathematical models describing the formation of beliefs. Model identification, data gathering, and data analysis become highlighted areas of current research due to the uncertainty about social learning mechanisms and networked interaction structure. To solve the identification problem, researchers carry out behavioral experiments and field investigations. In practice, researchers analyze communities on available real-world data, applying methods based on the structural properties of the network of information interactions between agents, the individual characteristics of agents, and a combination of structural and individual characteristics. Part III of the survey presents studies on identifying belief formation models and discusses some practical aspects of analyzing information communities in social networks.

Keywords: social networks, information community, formation of information communities, belief formation, detection of information communities.

INTRODUCTION

In parts I and II of the survey (see [1, 2]), the problems of identifying (detecting) and studying information communities in social networks were introduced. In addition, mathematical models of belief dynamics and the formation of information communities in social networks were presented, and the factors and conditions for the formation of information communities were considered. In practice, the identification of such models is nontrivial: many parameters are exogenous, and a significant aspect of the learning process remains unobservable in applied research. In many situations, people demonstrate neither their true beliefs nor information available for decision-making (the mechanisms for processing this information).

During social interaction, people receive incomplete information from their opponents, e.g., information about the results of actions (activity) of other people but not why and how these decisions were made. Many factors can cause this limitation, e.g., the nature of social interaction means or the high costs of receiving and (or) transmitting complete information. Researchers conduct field investigations and behavioral experiments to identify the real-life mechanisms of information processing by people despite the arising difficulties. Numerous methods for analyzing information communities were proposed using examples of publicly available data.

Part III of the survey is organized as follows. Section 1 discusses publications on the identification of belief formation models in networks. Section 2 considers applied research of information communities in social networks.

1. IDENTIFICATION OF BELIEF FORMATION MODELS IN NETWORKS

Depending on information processing methods, two types of agents can be distinguished (see parts I and II of the survey [1, 2]):

[1]This research was supported by the Russian Foundation for Basic Research, projects nos. 19-17-50225 and 20-07-00190 (D.A. Gubanov), and 20-31-90133 (I.V. Petrov).

- rational agents (e.g., within the concept of Bayesian rationality), which can further be divided into *myopic* agents (choosing the best response in the short term) and strategic agents (choosing an optimal response based on some game-theoretic concept, e.g., Nash equilibrium);
- *naive* agents, often described by DeGroot's rule (following it, the agents form their beliefs by averaging the observed opinions of other agents).

Early research works were devoted to identifying the types of agents in the laboratory and field investigations [3–5]. The existence of agents with different learning mechanisms within the same group was not assumed, which is a drawback of these studies. In the recent paper [6], an attempt was made to identify the types of agents on several sets of real data: it was shown that social groups consist of a mixture of rational (Bayesian) and naive (acting according to DeGroot's rule) agents, and the relationship between the types varies for different data sets. For example, in a series of behavioral experiments involving residents of 19 Indian villages, the identification procedure yielded the following results: 10% of the population behave in accordance with Bayesian rationality, and the rest of the agents prefer to average the responses of their neighbors in the social relations network. In the same experiment involving the students of the National Technological Institute of Mexico, the share of Bayesian agents reached 50%. The number of experiments for each group was 95 and 50, respectively, and the number of participants was 665 and 350, respectively.

The authors [6] identified four learning patterns to distinguish the agents with Bayesian and DeGroot's rule-based learning in the model with incomplete information. Moreover, they identified a key network characteristic separating learning types, called a *clan* (a strongly connected component of the graph):

1) If the clan consists entirely of agents following DeGroot's rule, who reach consensus on the state of the world at some time instant, then they will not change their ideas at the subsequent time instants (even if they are false).

2) In the model with complete information, Bayesian agent i, whose neighbors belong to the set of neighbors of Bayesian agent j, copies the estimate of the state of the world of agent j.

3) Regardless of the type of agent i, Bayesian agent j never considers his estimate of the state of the world (complement of pattern 2).

4) Even in the case of incomplete information, a Bayesian agent identifying the simple majority of the estimates of the state of the world of his neighbors will never change his estimate under any changes in the estimates of his particular neighbors.

In [7], a smaller-scale experimental study with similar results was carried out: the authors discovered that the agents' decisions agree with DeGroot's rule in 80–98% of cases, and the forecast errors depend on the agent's position in the network. However, the central observation was as follows: the real learning process of agents matches the naive DeGroot rule only in comparative statics, and the dynamics of reaching consensus in the laboratory experiments have more complex rules for information processing. Moreover, the authors identified these heuristics, tested a wide class of other learning rules, and modified the classical DeGroot model by allowing the agents to adjust the weight of their previous states.

In addition to the complexity of identifying information processing mechanisms by agents, the structure of social interaction is often difficult (or even impossible) to detect for an external observer. At the same time, this structure may have a crucial effect on agents' learning [8–10]. From this point of view, modern technologies (e.g., implemented via online social platforms) have radically changed the way people interact and consume information. Nevertheless, some phenomena preventing the identification of social relations and sources of information arise here as well. A key aspect in this area is the policy of processing personal data by online platforms [11], when a user has to choose between privacy and the disclosure of various personal information (biography, geolocation data, or the so-called digital traces—the history of activity on the Internet) to other users, owners of the platform, or third-party applications. Therefore, a user has to decide on the availability of information about his social relations with other network members. In addition, despite the increased efficiency of information transmission, users still have cognitive and temporal limitations. As a result, recommender systems have been developed, and there is an increasing interest in algorithmic personalization. The effect of algorithmic filtering on social learning is still underinvestigated, but several models (for example, see [12]) showed that the order of information messages received can significantly influence the effectiveness of learning and consensus reaching. All these factors play a decisive role in identifying the structures of information interaction and complicate the observability of social relations.

Thus, uncertainty arises both about the mechanisms of information processing by individuals and the structure of interactions within which agents exchange their information. These features motivate further research in the area.

2. STUDIES OF INFORMATION COMMUNITIES IN SOCIAL NETWORKS

2.1. Detection of Information Communities

There is no consensus in the literature regarding the formal definition of information communities. In applied research, the authors choose fairly general definitions that reasonably reflect the essence of the phenomena occurring in information interaction networks. Several such phenomena indicate the presence of information communities, together usually characterized as *controversy*:

– *echo chamber*, a socio-psychological phenomenon when opinions or beliefs are supported in communities of like-minded people approving and strengthening each other's opinions;

– *filter bubble*, a phenomenon when the algorithms of personalized recommender systems offer content consistent with the information earlier received by a user, thereby excluding his opportunity to get acquainted with alternative or new information.

The overwhelming majority of research into information communities is associated with significant restrictions. Such publications consider public opinions about political issues, focusing on large-scale and long-term events (e.g., elections). In many countries, citizens actively discuss socially significant issues on online social networks (Twitter, Facebook, etc.). As a result, huge thematic data sets containing information about users and their actions become available for analysis. Therefore, many works can be characterized as case studies, in which information communities are investigated on a specific data set related to a particular social phenomenon.

In these studies of information communities, as a rule, the processes of information propagation and their properties are considered; the formation mechanisms of network participants' beliefs are not identified. (For these problems, see the corresponding models in parts I and II of the survey [1, 2].) In addition to the complex identification procedure for learning rules, the reason is that in most theoretical models, the network structure is determined exogenously and does not depend on learning results: learning does not change the mutual influence of participants of the information process. However, empirical studies of the phenomena characterizing information communities reveal evidence of a relationship between learning and the structure of interactions. Identifying and formalizing these phenomena in theoretical models could significantly reduce the gap between theory and practice. Modern applied research is limited to the development of identification methods for the state of individuals (assessment of private beliefs based on observed information) and the analysis of comparative statics.

Attempts to identify the states of information interaction participants often rest on the following observation. Generally, the formation of an information community can be represented as a diffusion process on a network (known as the diffusion of innovations, ideas, or information) in which joining a new community is analogous to the acceptance of ideas or beliefs. The converse is also true: any propagation process on a network can be viewed as the formation of a network community in which the elements are grouped by their states. One example of such processes is information propagation called information cascade; see Fig. 1.

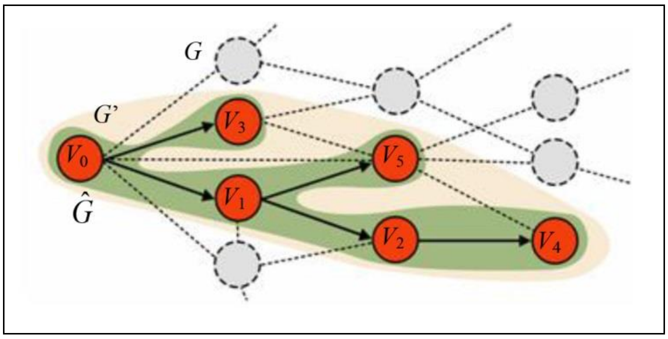


Fig. 1. Community formation process interpreted as a diffusion process on the network [13]. The thick directed edges in the graph G show the information propagation process starting in vertex V_0 and covering the vertices of the subgraph $\hat{G}$. The undirected edges show the relations of social interactions between network nodes (e.g., friendship ties). All together, these relations induce a friends subgraph G'.

According to a natural assumption, such information processes (cascades) should correlate with the beliefs of the individuals involved and affect their beliefs. This analogy with diffusion processes often becomes a starting point when studying the formation of information communities: the authors apply methods based on the structural properties of the information interaction network, the properties of the network elements, or a combination of structural and individual characteristics.

Thus, the problem is identifying two main characteristics of information interaction: the communication structure of the participants and their individual characteristics. The table below presents the most cited papers on identifying information communities: a brief description of the data used, the proposed measures, and the methods adopted by the authors. As mentioned above, both identification problems are complex, and the choice of an appropriate method for identifying information communities largely depends on the set of real data available to the researchers.

Most cited papers on identifying information communities

Paper title	Concept of community	Graph type	Measure (characteristic) type	Data type	Data source
Testing Models of Social Learning on Networks: Evidence from Two Experiments [6]	A set of nodes more connected to each other than to those outside the group	An artificially created network of relations between the experiment participants	Structural (clan)	Offline	Laboratory experiments
Ideological Segregation Online and Offline [14]	A community with equivalent characteristics of members	Relations between interaction participants are not considered	Individual characteristics (isolation index)	Online, offline	Internet news, offline media, personal interaction
Quantifying Controversy in Social Media – [15]	Opinions or beliefs are supported in the communities created by like-minded people, who strengthen and approve of each other's opinions	A dialog graph: a graph corresponding to thematic discussions, where relations between participants are formed in the case of users' responses to each other's messages	Structural (Random Walk Controversy, Betweenness Centrality Controversy, Embedding Controversy)	Online	Twitter
Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship [16]	Preferences for the content received by network users match the preferences for the content they distribute	A subscriber graph: a directed edge (link) between participants arises if one participant monitors information updates from another participant	Individual characteristics (production polarity, consumption polarity)	Online	Twitter
Community Interaction and Conflict on the Web [17]	Community members interact primarily with other members of their community	A bipartite multigraph between users and communities. Relations arise in the case of communication between users within a given community	Mixed	Online, time series	Reddit
Quantifying Echo Chamber Effects in Information Spreading over Political Communication Networks [18]	Beliefs are strengthened through repeated interactions with people sharing the same viewpoints	A subscriber graph: a directed edge (link) between participants arises if one participant monitors information updates from another participant	Mixed	Online, time series	Twitter
An Empirical Examination of Echo Chambers in US Climate Policy Networks [19]	A community is characterized by two attributes: information coinciding with established beliefs and a clustered structure of interaction	A network of interaction between experts	Mixed	Online	Survey
Echo Chambers: Emotional Contagion and Group Polarization on Facebook [20]	Groups of like-minded people with extreme-value beliefs	Relations between interaction participants are not considered	Individual characteristics (user sentiment polarization)	Online	Facebook
Exposure to Ideologically Diverse News and Opinion on Facebook [21]	Two types of communities: – a set of participants exposed only to information from like-minded people; – the information offered by the algorithms matches the history of user actions	A graph of friendship ties between social network participants	Mixed (Alignment score)	Online	Facebook
Filter Bubbles, Echo Chambers, and Online News Consumption [22]	Two types of communities: – a set of participants exposed only to information from like-minded people; – the information offered by the algorithms matches the history of user actions	Relations between interaction participants are not considered	Individual characteristics (audience-based measure of outlet slant)	Online	Net-surfing history

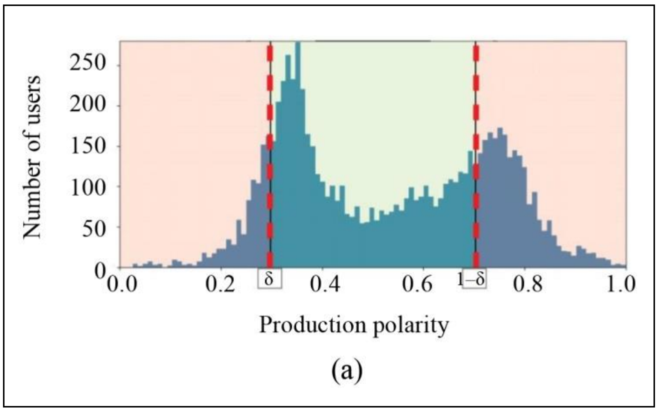
Subsection 2.2 provides some of the most common methods and measures for identifying information communities.

2.2. Identification of Information Communities: Methods Based on the Properties of Network Structure Elements

In applied research, content is crucial for measuring the effects that characterize the presence of information communities. The essential characteristic of an information community is the degree of correspondence between the content consumed and produced by the network participants. In this regard, the authors [16] divided the general process of information interaction into information consumption and information production processes. According to their approach, each message t of a social network belongs to one of two subclasses: $l(t)=l_n\in\{0,1\}$. (The study involved data from Twitter with the following notations: P_u is information published by the user on his page; C_u is publications that the user receives from other users. The information was classified by the users' political views, where l_n = 1 for conservative views and l_n = 0 for the liberal ones.) Based on the set of all information produced (P_u) and consumed (C_u) by the user u, the degree of diversity of the content produced and consumed by the users (production and consumption polarity, respectively) was determined as the amount of information from one class divided by the total amount of information produced and consumed by the user:

$$p(u)=\frac{\sum_{t\in P_u} l(t)}{|P_u|}\ ,\quad c(u)=\frac{\sum_{t\in C_u} l(t)}{|C_u|}\ .$$

There are variations of these measures (where the variability is due to the specifics of real data), yielding practically interpretable conclusions about information propagation among the users. In particular, by calculating statistical characteristics (variance, correlation, or measures of difference between probability distributions) for the diversity of consumed and produced content, the researchers demonstrated the presence of information communities; see Fig. 2.



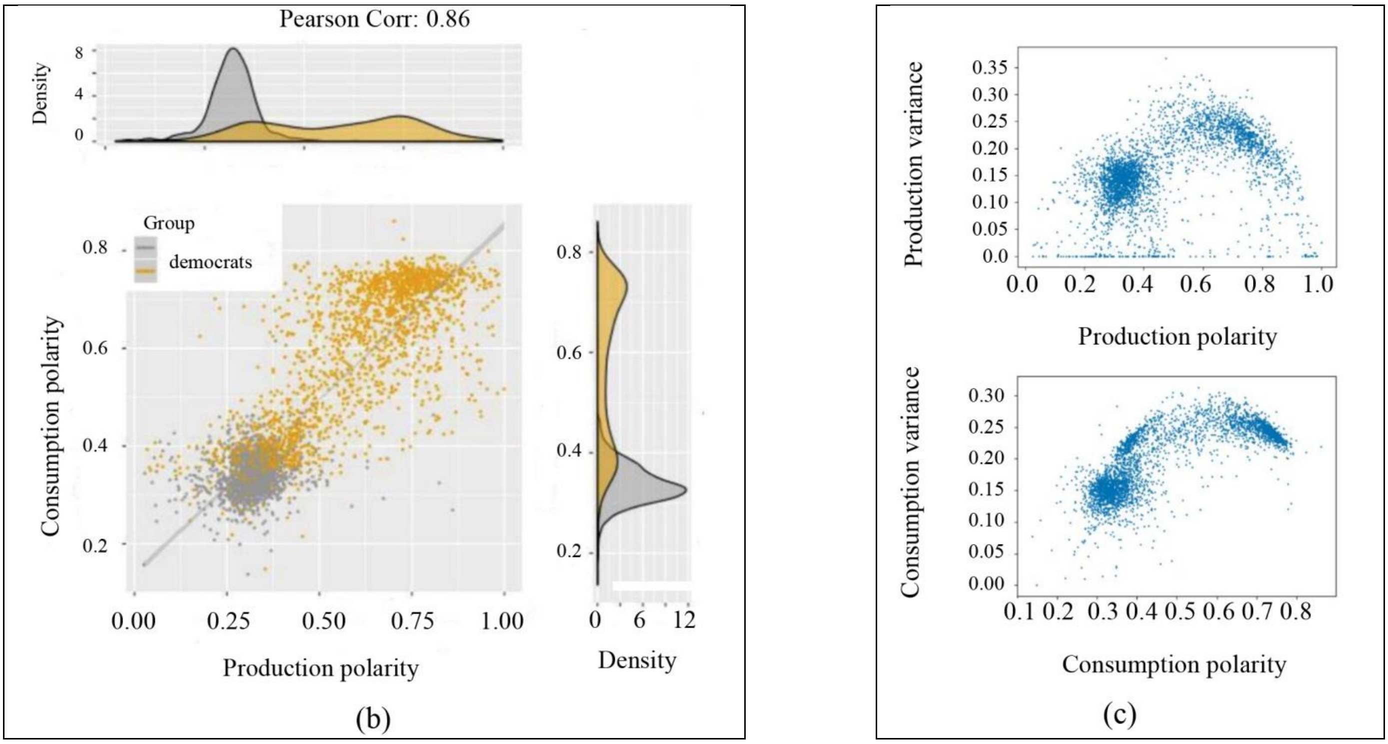


Fig. 2. Estimated statistical characteristics of information interaction between Twitter users on the legislative regulation of arms traffic [16].

Figure 2*a* presents the distribution of users by production polarity: the double-peaked property of this distribution indicates, among other signs, the presence of information communities (echo chambers). The graph of the relationship between consumption and production polarity (Fig. 2*b*) demonstrates the high-degree clustering of the values for representatives of different user groups. The graphs on the right are intended to assess the relationship between production (consumption) polarity and its variance.

2.3. Analysis of Information Communities: Methods Based on the Structural Properties of Information Interaction

Analyzing the structural properties of information communities, researchers focus on comparing interaction processes between the nodes of different communities. The analysis tool is often the properties of random processes on graphs or centrality measures reflecting the effectiveness of nodes during information propagation.

Random Walk Controversy (RWC [19]) is a random walk-based measure defined as follows. Let the graph be partitioned by some criterion into two subgraphs, *X* and *Y*, with nonintersecting sets of vertices. Consider two random walks, one ending in the subgraph *X* and the other in the subgraph *Y*. RWC is the difference between the probabilities of two events: (1) both random walks started in the same subgraph where they ended and (2) both random walks started in a subgraph differing from the one where they ended). That is,

$$\mathrm{RWC} = P_{XX} P_{YY} - P_{YX} P_{XY},$$

where

$P_{AB} = P(\text{the process starts in } A \mid \text{the process ends in } B)$

denotes the corresponding conditional probability with $A, B \in \{X, Y\}$.

Another random walk-based method for identifying information communities is a personalized version of the PageRank algorithm [23], in which the damping factor changes depending on the group of the graph vertex in which the random walk process starts [17]. In the classical version of the algorithm, transitions occur either to neighbors or any other node selected equiprobably. (A practical interpretation is the end of the link click process and the beginning of a new one). In the personalized version, the probability distribution on the set of vertices is different for vertices from different communities. Thus, the method allows assessing the controversy of communities by comparing the probabilities of interaction between members of different communities [17].

A measure based on betweenness centrality [15]. The betweenness centrality $bc(e)$ of a network edge e is defined as

$$bc(e) = \sum_{s \neq t \in V} \frac{\sigma_{s,t}(e)}{\sigma_{s,t}},$$

where $\sigma_{s,t}$ is the total number of shortest paths between vertices s and t in the graph, and $\sigma_{s,t}(e)$ is the number of shortest paths passing through the edge e. The authors [15] proposed to analyze the differences in the centralities of the vertices from two sets forming a graph partition. (In the original work, they used the METIS algorithm [24].) The idea is to compare the centralities of the edges included in the graph cut-set (i.e., the edges connecting the vertices from different subsets of the graph vertices) and the centralities of the edges in the rest of the graph. If a "good" graph cut-set is obtained, most of the shortest paths from one graph part to another will pass through the cut-set edges, and the centrality of these edges will have higher values compared to the centrality of the edges in the rest of the graph. Comparing two distributions of centralities––inside the cut-set and outside it—for example, using the KL-divergence d_{KL} and performing normalization, we obtain the following expression for *Betweenness Centrality Controversy* (BCC):

$$\mathrm{BCC} = 1 - e^{-d_{KL}}.$$

In addition to these methods, classical clustering techniques without considering diffusion processes on networks or calculating paths between vertices are used. Researchers associate the resulting structural characteristics with the individual characteristics of separate graph nodes, thereby combining the methods demonstrated above.

2.4. Analysis of Information Communities: Methods Based on the Combination of Structural and Individual Characteristics

Combining the individual characteristics of the participants and the structural characteristics of information interaction is a nontrivial problem underinvestigated in the literature. One solution is to employ machine learning methods: all characteristics of the information interaction process available to researchers (the set of all produced or consumed information, information content, structural characteristics of the network and individual participants, etc.) are considered and placed in a single feature space. Here, classification methods are used to detect information communities.

When combining individual and structural characteristics, a promising line is to adopt various transformations of the initial data: *embedding* and, particularly, node/edge/graph embedding [25].

This operation generally transforms the original feature space into another space, often of a lower dimension. From this viewpoint, all the methods mentioned above can be understood as special cases of such transformations. The clustering problem can be solved by classical methods in a new space [26, 27]. For example, the *Embedding Controversy* (EC) measure

$$\mathrm{EC} = 1 - \frac{d_X + d_Y}{2d_{XY}},$$

where d_X (d_Y) is the average distance between the pairs of elements from the set X (Y, respectively), and d_{XY} is the average distance between the pairs of elements from different sets, yields another method for identifying information communities [15]. An EC value close to 1 indicates the presence of information communities and a high degree of graph clustering; an EC value close to 0 indicates the opposite.

The graph embedding method is more complicated; however, it allows analyzing not individual nodes but entire graphs. The method involves graph kernels—transformations for the pairwise comparison of structures with each other—and can be used both for comparative analysis of individual groups of graph vertices [28] and information processes occurring on networks [29–31]. This approach allows studying the sequences of information flows and comparing and predicting the characteristics of information cascades (such as size, speed, etc.) in information communities.

CONCLUSIONS

This paper has overviewed studies of information communities in complementary areas: the formation models of information communities in social networks (with microeconomic, cognitive, and socio-psychological foundations), identification methods for information communities, and applied research into information communities in social networks.

Parts I and II of the survey have outlined the concept of an information community and considered belief formation models for individuals seeking to eliminate uncertainty about (a) given question(s), eventually forming information communities. Approaches to model the belief updating process of individuals and the effect of various factors on reaching true beliefs and forming different (or identical) stable beliefs in the network, leading to the emergence of information communities, have been described.

This part of the survey has presented empirical studies of the existence of information communities in real social networks. Uncertain mechanisms of information processing by individuals, an uncertain structure of interaction, and abundant real data sets (mainly from online social networks) cause a wide variety of empirical methods for identifying information communities and research focusing on real data sources. Due to the specifics of the available data, the considered methods characterize the information produced and consumed by social network users rather than their beliefs. The absence of any prerequisites for belief formation mechanisms is a significant drawback of these methods: only indirect conclusions can be drawn both about the true beliefs of the participants in information interactions and the formation dynamics of information communities. The transition from the analysis of individual node interactions to the analysis of higher-order structure interactions characterizing the evolution of the information process seems promising for identifying communities in information interaction structures. Investigations in this area can significantly expand the understanding of the relationship between information processes and the formation of information communities.

REFERENCES

1. Gubanov, D.A. and Petrov, I.V., Information Communities in Social Networks. Part I: From Concept to Mathematical Models, *Control Sciences*, 2021, no. 1, pp. 13–20.
2. Gubanov, D.A. and Petrov, I.V., Information Communities in Social Networks. Part II: Networked Models of Formation, *Control Sciences*, 2021, no. 2, pp. 16–28.
3. Hung, A.A. and Plott, C.R., Information Cascades: Replication and an Extension to Majority Rule and Conformity-Rewarding Institutions, *American Economic Review*, 2001, vol. 91, no. 5, pp. 1508–1520.
4. Kübler, D. and Weizsäcker, G., Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory, *The Review of Economic Studies*, 2004, vol. 71, no. 2, pp. 425–441.
5. Choi, S., Gale, D., and Kariv, S., *Behavioral Aspects of Learning in Social Networks: An Experimental Study*, Emerald Group Publishing Limited, 2005.
6. Chandrasekhar, A.G., Larreguy, H., and Xandri, J.P., Testing Models of Social Learning on Networks: Evidence From Two Experiments, *Econometrica*, 2020, no. 1 (88), pp. 1–32.
7. Grimm, V. and Mengel, F., Experiments on Belief Formation in Networks, *Journal of the European Economic Association*, 2020, vol. 18, no. 1, pp. 49–82.
8. Acemoglu, D. and Ozdaglar, A., Opinion Dynamics and Learning in Social Networks, *Dynamic Games and Applications*, 2011, no. 1 (1), pp. 3–49.
9. Golub, B. and Sadler, E., Learning in Social Networks, in *The Oxford Handbook of the Economics of Networks*, Bramoullé, Y., Galeotti, A., and Rogers, B.W., Eds., 2016. DOI: 10.1093/oxfordhb/9780199948277.013.12.
10. Young, J.G., Cantwell, G.T., and Newman, M.E. J., Robust Bayesian Inference of Network Structure from Unreliable Data, *arXiv preprint arXiv:2008.03334*, 2020.

11. Lovato, J., Allard, A., Harp, R., and Hébert-Dufresne, L., Distributed Consent and Its Impact on Privacy and Observability in Social Networks, *arXiv preprint arXiv:2006.16140*, 2020.
12. Perra, N. and Rocha, L.E.C., Modelling Opinion Dynamics in the Age of Algorithmic Personalisation, *Scientific Reports*, 2019, vol. 9, no. 1, pp. 1–11.
13. Cheng, J., Adamic, L., Dow, P.A., et al., Can Cascades Be Predicted? *Proceedings of the 23rd International Conference on World Wide Web*, 2014, pp. 925–936.
14. Gentzkow, M. and Shapiro, J.M., Ideological Segregation Online and Offline, *The Quarterly Journal of Economics*, 2011, vol. 126, no. 4, pp. 1799–1839.
15. Garimella, K., Morales, G.D.F., Gionis, A., and Mathioudakis, M., Quantifying Controversy on Social Media, *ACM Transactions on Social Computing*, 2018, vol. 1, no. 1, pp. 1–27.
16. Garimella, K., Morales, G.D.F., Gionis, A., and Mathioudakis, M., Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship, *ArXiv180101665 Cs*, 2018.
17. Kumar, S., Hamilton, W.L., Leskovec, J., and Jurafsky, D., Community Interaction and Conflict on the Web, *Proceedings of the 2018 World Wide Web Conference*, 2018, pp. 933–943.
18. Cota, W., Ferreira, S.C., Pastor-Satorras, R., and Starnini, M., Quantifying Echo Chamber Effects in Information Spreading over Political Communication Networks, *EPJ Data Science*, 2019, vol. 8, no. 1, p. 35.
19. Jasny, L., Waggle, J., and Fisher, D.R., An Empirical Examination of Echo Chambers in US Climate Policy Networks, *Nature Climate Change*, 2015, vol. 5, no. 8, pp. 782–786.
20. Del Vicario, M., Vivaldo, G., Bessi, A., et al., Echo Chambers: Emotional Contagion and Group Polarization on Facebook, *Scientific Reports*, 2016, vol. 6, p. 37825.
21. Bakshy, E., Messing, S., and Adamic, L.A., Exposure to Ideologically Diverse News and Opinion on Facebook, *Science*, 2015, vol. 348, no. 6239, pp. 1130–1132.
22. Flaxman, S., Goel, S., and Rao, J.M., Filter Bubbles, Echo Chambers, and Online News Consumption, *Public Opinion Quarterly*, 2016, vol. 80, no. S1, pp. 298–320.
23. Page, L., Brin, S., Motwani, R., and Winograd, T., The PageRank Citation Ranking: Bringing Order to the Web, Stanford InfoLab, 1999.
24. Karypis, G., METIS: Unstructured Graph Partitioning and Sparse Matrix Ordering System, Technical Report, 1997.
25. Rossi, R.A., Jin, D., Kim, S., et al., On Proximity and Structural Role-Based Embeddings in Networks: Misconceptions, Techniques, and Applications, *ACM Transactions on Knowledge Discovery from Data* (*TKDD*), 2020, vol. 14, no. 5, pp. 1–37.
26. Chattopadhyay, S. and Ganguly, D., Community Structure Aware Embedding of Nodes in a Network, *arXiv preprint arXiv:2006.15313*, 2020.
27. Billings, J.C.W., Hu, M., Lerda, G., et al., Simplex2Vec Embeddings for Community Detection in Simplicial Complexes, *arXiv preprint arXiv:1906.09068*, 2019.
28. Rossi, R.A., Jin, D., Kim, S., et al., From Community to Role-Based Graph Embeddings, *arXiv preprint arXiv:1908.08572*, 2019.
29. Zhou, X. and Zafarani, R., A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities, *arXiv preprint arXiv:1812.00315*, 2018.
30. Ma, J., Gao, W., and Wong, K.F., Detect Rumors in Microblog Posts Using Propagation Structure via Kernel Learning, Association for Computational Linguistics, 2017.
31. Wu, K., Yang, S., and Zhu, K.Q., False Rumors Detection on Sina Weibo by Propagation Structures, *2015 IEEE 31st International Conference on Data Engineering*, 2015, pp. 651–662.

This paper was recommended for publication by D.A. Novikov, a member of the Editorial Board.

Received August 26, 2020, and revised November 9, 2020.
Accepted November 24, 2020.

Author information

Boiko, Liliya Mikhailovna. Junior researcher, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ boiko.lilia@gmail.com

Gubanov, Dmitry Alekseevich. Cand. Sci. (Eng.), Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ dmitry.a.g@gmail.com

Petrov, Il'ya Vladimirovich. Postgraduate, Trapeznikov Institute of Control Sciences, Russian Academy of Sciences, Moscow, Russia, ✉ zyxzy@protonmail.ch

Cite this article

Boiko, L.M., Gubanov, D.A., Petrov, I.V. Information Communities in Social Networks. Part III: Applied Aspects of Detection and Analysis. *Control Sciences* **3,** 14–21 (2021).
http://doi.org/10.25728/cs.2021.3.2