On the Concept of an Informational Community in a Social Network

A G Chkhartishvili, D A Gubanov
V.A. Trapeznikov Institute of Control Sciences of RAS, Moscow, Russia
dmitry.a.g@gmail.com

Abstract. This paper introduces a constructive definition of an informational community, which agrees with formal models of opinion dynamics for bounded rational agents in social networks. As is shown below, uncertainty can be taken into account when detecting informational communities. An example of a stable informational community is given. A control problem for informational communities is stated, and its solution is presented within the DeGroot model.

1. Introduction
In many subject areas, an important practical problem is to detect and study informational communities in online social networks. An informational community is a set of individuals with common and stable opinions about some issue. For example, in sociology and political science, the formation of isolated communities is often supposed to cause polarization and conflicts. Theoretical and applied research deals with the issues of obtaining alternative information by a user depending on the preferences of his/her contacts and online social network algorithms [1], the issues of interaction between communities with different opinions [2], the issues of informational roles of users [3], etc. Statistical and machine learning methods, as well as methods of matrix analysis and social network analysis [1–5], are often used. The features of information processing by an individual are studied in cognitive science, psychology, and social psychology (for example, see [6]), and formal models of bounded rational agents are developed to describe opinion dynamics in networks, taking into account these features. The classical model was suggested by DeGroot [7]: in this model, an agent updates his/her own opinion based on the available information about the opinions of his/her network environment. In models closer to reality [8–12], the degree of influence of agent’s neighbors depends on how much their opinions agree with his/her one.

In practice, it is difficult to use these models for detecting communities, due to the simplifications and assumptions accepted in the course of modeling, due to the complexity of identifying model parameters, and, finally, due to the fact that models of opinion dynamics are not enough for considering informational communities. This paper proposes a constructive definition of an informational community based on the conceptual and formal models of opinion dynamics in social networks. A control problem for informational communities is stated, and its solution is presented within the DeGroot model.

2. Informational community

2.1. Concept of informational community.
Consider a directed network of informational influence, in which the nodes are agents from the set \( N = \{1, \ldots, n\} \). For each agent \( i \), denote by \( N_i \) the set of nodes influencing him/her. Assume that the network has a fixed structure.
Each agent is characterized by his/her state \( x_i \in S \), which reflects his/her opinion about some issue. In a practical interpretation, the agent’s state can mean his/her awareness of some problem, a complex of his/her ideological and political views, etc. The agent’s state is expressed through the actions he/she performs; for example, subscribing to some pages on the network, writing posts of definite content, etc.

In the sequel, the accounts of ordinary individuals in online social networks will be distinguished from the accounts of groups, public pages, etc. Assume that changes in an agent’s state are influenced by his/her environment composed of similar agents, which can also change their states. As for the pages to which an agent is subscribed, they only maintain his/her current state, also being state indicators.

The agent’s state evolves in discrete time depending on the environment influencing him/her:

\[
x_i(t+1) = F_i(x_i(t), x_{-i}(t), \theta),
\]

where \( x_{-i}(t) \) denotes the state vector of agents from the set \( N_i \) at step \( t \); \( \theta \in \Theta \) is an unknown parameter. (In other words, the dynamics of the state vector are known up to this parameter.)

Consider the problem of detecting informational communities from the viewpoint of a control authority (further referred to as the Principal). At some initial step \( t_0 \), being guided by his/her own interests and a priori beliefs, the Principal determines the planning horizon \( k \) (a natural number) and the sets of states \( X \subseteq S \) and agents \( M \subseteq N \) of interest to him/her.

**Definition.** A subset \( M \subseteq N \) of agents is called an informational community if the states of all agents from \( M \) belong to the set \( X \) at all steps \( t_0, t_0+1, ..., t_0+k \), i.e.,

\[
\forall i \in M, \forall \theta \in \Theta, \forall t \in \{t_0, t_0+1, ..., t_0+k\}: x_i(t) \in X.
\]

In a practical interpretation, this condition means that the information received by the agents does not cause (at the next \( k \) steps) a significant revision of their opinions for any values of the unknown parameter, and close opinions hold in the community. It can be said that all members of the community are like-minded, and their common views remain stable.

In applications, the parameter \( k \) is the planning horizon on which the network dynamics are considered. (Recall that the network is assumed to be constant.) If slowly changing states are the subject of study, one should consider informational communities for small values of the parameter \( k \), e.g., 1 or 2.

**Example** (the DeGroot model). Let \( k = 1 \) and the state dynamics of agent \( i \) be defined as a convex combination of his/her state and the state of his/her environment:

\[
x_i(t+1) = (1 - \theta_i)x_i(t) + \theta_i \sum_{j \in N_i} b_{ij} x_j(t),
\]

where the coefficients \( b_{ij} \) specify the relative influence exerted on agent \( i \) by agent \( j \) from the former’s environment \( N_i \); the coefficient \( \theta_i \) can be treated as the degree of influence exerted on agent \( i \) by the entire environment. (All these coefficients take values within the range \([0, 1]\).) When analyzing data of a real online social network, the coefficients \( b_{ij} \) can be calculated, e.g., using the actional model of influence [13, 14]. At the same time, it can be difficult to estimate the coefficients \( \theta_i \). Hence, a dynamic model with uncertainty of the form (1) arises accordingly. Applying condition (2) gives

\[
\forall i \in M, \forall \theta_i \in [0, 1]: (1 - \theta_i)x_i(t_0) + \theta_i \sum_{j \in N_i} b_{ij} x_j(t_0) \in X.
\]

Clearly, in the case of a convex set \( X \) (which seems quite natural in applications), this condition can be written as

\[
\forall i \in M: x_i(t_0) \in X, \sum_{j \in V_i} b_{ij} x_j(t_0) \in X.
\]

This means that the states of all agents from \( M \) belong to the set \( X \) and, in addition, the influence of the environment on each agent does not cause his/her state to leave the set \( X \).
Consider an example of an informational community for a social network demonstrated in Fig. 1a. (Here the numbers inside network nodes indicate the numbers of agents.) Let $\theta_i = 0.1$ for $i = 5,15$ and $\theta_i = 0.5$ for $i = 1,4$. The initial opinions of agents belong to the range $S = [0,1]$. (In Fig. 1a, the stronger the node’s color is, the smaller opinion the corresponding agent will have.) The Principal fixes the pole of opinions $X = [0.9,1.0]$ of interest to him/her, as well as chooses the step $t_0 = 4$ and the planning horizon $k = 1$. Then the Principal determines a potential community $M = \{i \in N|x_i(t_0) \in X\}$. According to Fig. 1b, the set $M$ is an informational community.

![a) social network](image1.png) ![b) trajectories of agents’ opinions](image2.png)

**Figure 1. Example of informational community.**

2.2. Stable informational communities

Like before, let the Principal fix some set of states $X \subseteq S$ and some initial step $t_0$.

**Definition.** A subset $M \subseteq N$ of agents is called a stable informational community if it remains an informational community for any $k \in \{0,1,2,...\}$. This means that the states of all agents from $M$ will stay within the set $X$ even on large planning horizons (under the existing assumptions regarding the state dynamics).

**Example.** Consider an informal example of a stable informational community as follows. Let the set $X$ be finite, and the state of agent $i$ at each step be determined by the state of the majority of agents in his/her influencing environment. Also, assume that at a step $t_0$ there exists a subset $M \subseteq N$ of agents such that: (1) all agents from this subset are in the same state $x \in X$ and (2) for each agent $i \in M$, the majority of his/her influencing environment $V_i$ consists of the like-minded agents from $M$. Then $M$ is a stable informational community.

Now consider an example of a stable informational community for the social network of Florentine families [16], shown in Fig. 2a. Let $S = \{0,1\}$, and note that all nodes with the initial state 0 are colored white in Fig. 2a. The Principal fixes $X = \{0\}$ and $t_0 = 4$. In this case, the set $M = \{Bischeri, Castellani, Peruzzi, Strozzi\}$ is a stable informational community; see Fig. 2b.

![a) social network](image3.png) ![b) dynamics of the number of agents](image4.png)

**Figure 2. Example of stable informational community.**
2.3. Control of informational communities

A problem of informational control [10] is the control of informational communities. Consider the case in which a control authority (the Principal) seeks to minimize the number of agents in an informational community at the next step (i.e., for \( k = 1 \)), influencing their current opinions and, thereby, the opinion dynamics (1). In a practical interpretation, this influence consists, e.g., in the proposal to subscribe to the pages expressing alternative points of view, in the application of an appropriate algorithm for forming the user’s news feed, etc.

Then a control problem can be stated in the following way:

\[
\Phi(u) = |\{i \in N | F_i(t_0), x_{i\ldots(i)}(t_0), \theta, u \in X\}| \quad \text{min.} \quad \text{subject to } u \in U,
\]

where \( \Phi(u) \) is the Principal’s goal function; \(|\cdot|\) denotes the set cardinality; \( \theta \) is a fixed and known parameter; \( U \) is the set of admissible informational controls of the Principal. The solution of the problem (3) is an optimal informational influence of the Principal in the situation described above.

Example. Let the dynamics of agents’ opinions be given by the DeGroot model, and the control parameter be the degree of influence \( u \in U = [0, 1] \) exerted on each agent by his/her environment, the same for all agents:

\[
x_i(t + 1) = (1 - u)x_i(t) + u \sum_{j \in N_i} b_{ij} x_j(t).
\]

In addition, let \( S = [0, 1] \) and \( X = \{x', x''\} \subset S \), i.e., the set of all undesired states, from the Principal’s viewpoint, is a closed interval.

Introduce the notation

\[
y_i(t) = \sum_{j \in N_i} b_{ij} x_j(t).
\]

Then the problem (3) takes the form

\[
\Phi(u) = |\{i \in N |(1 - u)x_i(t_0) + uy_i(t_0) \in X\}| \quad \text{min.} \quad \text{subject to } u \in U.
\]

As is easily checked, for each agent \( i \in N \), the set of all values \( u \) satisfying the inclusion

\[
(1 - u)x_i(t_0) + uy_i(t_0) \in X,
\]

is either empty (in the cases \( x_i(t_0) < x' \), \( y_i(t_0) < x' \) and \( x_i(t_0) > x'' \), \( y_i(t_0) > x'' \) or represents a closed interval (possibly, degenerating into a point). Denote this interval by \([u'_i, u''_i]\). (Recall that the cases \( u'_i = u''_i, u'_i = 0, u''_i = 1 \) are possible.) For example, for \( x' \leq x_i(t) \leq x'' < y_i(t) \), the boundary points of the interval take the values \( u'_i = 0 \) and \( u''_i = (x'' - x_i(t))/(y_i(t) - x_i(t)) \), where \( u''_i \) is the solution of the equation \((1 - u)x_i(t) + uy_i(t) = x''\).

The Principal’s goal function \( \Phi(u) \) is piecewise constant, i.e., changes its values at the boundary points only. Arrange the values \( u'_i, u''_i \) for different \( i \) in the ascending order. (The total number of such values does not exceed \( 2n \).) Denote by \( \bar{U} \) the set containing all these values, the midpoints for the adjacent elements in the ordering, and the values 0 and 1. (The number of elements in the set \( \bar{U} \) does not exceed \((4n + 1)\).) Then the problem (4) can be written as

\[
\Phi(u') = |\{i \in N |(1 - u)x_i(t) + uy_i(t) \in X\}| \quad \text{min.} \quad \text{subject to } u \in \bar{U}.
\]

Note that the exhaustive enumeration of at most \((4n + 1)\) values of the function \( \Phi(u') \) at the points of the same set \( \bar{U} \) can be used to minimize (5) as well as to maximize the number of agents in an informational community.

3. Conclusions

In this paper, a new constructive definition of an informational community has been introduced. On the one hand, this definition agrees with formal models of opinion dynamics for bounded rational agents in social networks; on the other, it can be used to detect informational communities in real
online social networks. The examples of models with discrete and continuous opinions of agents have been considered to study the detection of informational communities under uncertainty and the existence of stable informational communities. A control problem for informational communities has been stated, and its solution has been presented within the DeGroot model.

Acknowledgments
This work was supported by the Russian Foundation for Basic Research, project number 18-29-22042.

References
[1] Bakshy E, Messing S and Adamic L A 2015 Exposure to ideologically diverse news and opinion on Facebook Science 348 1130–1132
[2] Garimella K, Morales G D F, Gionis A and Mathioudakis M 2018 Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship ArXiv180101665 Cs
[3] Kumar S, Hamilton W L, Leskovec J and Jurafsky D 2018 Community Interaction and Conflict on the Web Proc. of the 2018 World Wide Web Conf. World Wide Web pp. 933–943
[4] Barberá P, Jost J T, Nagler J, Tucker J A and Bonneau R 2015 Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? Psychol. Sci. 26 1531–1542
[5] Byzov L G, Chkhartishvili A G, Gubanov D A and Kozitsin I V 2020 Perfect politician for social network: an approach to analysis of ideological preferences of users Control Sciences Issue 4 15-26
[6] Myers D G 2010 Social psychology (New York, NY: McGraw-Hill)
[7] DeGroot M H 1974 Reaching a Consensus J. Am. Stat. Assoc. 69 118–21
[8] Deffuant G, Neau D, Amblard F and Weisbuch G 2000 Mixing Beliefs among Interacting Agents Advances in Complex Systems 03 87-98
[9] Hegselmann R and Krause U 2015 Opinion Dynamics under the Influence of Radical Groups, Charismatic Leaders, and Other Constant Signals: A Simple Unifying Model Networks and Heterogeneous Media 10 477-509
[10] Chkhartishvili A G, Gubanov D A and Novikov D A 2019 Social Networks: Models of Information Influence, Control and Confrontation 189 (Cham: Springer International Publishing)
[11] Novikov D 2020 Dynamics models of mental and behavioral components of activity in collective decision-making Large-Scale Systems Control Issue 85 206-237 (In Russian)
[12] Chkhartishvili A G, Gubanov D A and Petrov I V 2020 Multidimensional model of opinion dynamics in social networks: polarization indices Control Sciences Issue 3 26-33(In Russian)
[13] Gubanov D A and Chkhartishvili A G 2018 Influence Levels of Users and Meta-Users of a Social Network Autom. Remote Control 79 545–553
[14] Gubanov D 2020 A Study of Formalizations of User Influence in Actional Model Proc. of the 13th Int. Conf. "Management of Large-Scale System Development" (Moscow, Russia: IEEE)
[15] Gubanov D and Petrov I 2019 Multidimensional Model of Opinion Polarization in Social Networks Proc. of the 12th Int. Conf. “Management of large-scale system development” (MLSD) (Moscow, Russia: IEEE)
[16] Padgett J F and Ansell C K 1993 Robust Action and the Rise of the Medici, 1400-1434 America Journal of Sociology 98 1259-1319.