UDC 519.87 Вестник СПбГУ. Прикладная математика. Информатика... 2024. Т. 20. Вып. 2
MSC 90B15, 90B18, 90C40

Analysis of consensus time and winning rate in two-layer networks with hypocrisy of different structures*

Ch. Zhao, E. M. Parilina

St. Petersburg State University, 7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation

For citation: Zhao Ch., Parilina E. M. Analysis of consensus time and winning rate in two-layer networks with hypocrisy of different structures. *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*, 2024, vol. 20, iss. 2, pp. 170–192.
https://doi.org/10.21638/spbu10.2024.204

We have developed a microscopic version of general concealed voter model (GCVM). Original GCVM uses only statistical-physical methods, while our new approach starts with a real network. A microscopic model is suitable for any two-layer network (with internal and external layers) satisfying the definition given in the paper. We conduct a series of simulations with different network structures and found that a cyclic external structure prolongs consensus time in comparison with a complete external structure. Moreover, a cyclic external structure has a positive impact on a winning rate, and this result is different from the one obtained in the macroscopic version of GCVM. The possible reasons for this difference are discussed in the paper. Additionally, we propose and validate the hypothesis that there exists a strong linear relationship between a consensus time and pairwise average shortest paths d in the network structure. We performed a controlled variable approach to validate the impact of each individual parameter on key performance indicators (KPIs) including a consensus time and winning rate. Furthermore, we assess the influence of parameter combinations on KPIs by analyzing the results using the K-means algorithm. We conclude that certain parameter combinations can have a significant impact on the consensus time.

Keywords: opinion dynamics, voter model, concealed voter model, general concealed voter model, winning rate.

1. Introduction. Opinion dynamics models can be divided into two main groups: macroscopic and microscopic. Macroscopic models examine social networks using statistical-physical methods and applying probability theory and statistical methods to analyze the evolution of opinion distribution, e.g., the Ising model [1], voter model [2], concealed voter model (CVM) [3, 4], and a macroscopic version of the general concealed voter model (GCVM) [5]. The Ising model has a long history in statistical physics [6]. The Sznajd model [7] is one of the well-known modifications of the Ising model. In each round of the Sznajd model, a pair of agents a_i and a_{i+1} is selected for interaction to influence the nearest neighbors, i.e. agents a_{i-1} and a_{i+2}. In a voter model [2], a random agent a_i is chosen, then her random neighbor is chosen, and this neighbor adopts a_i's opinion. In CVM [3, 4], it is supposed that the social network is divided into external and internal layers, and the individuals conceal or publicly express their opinions. The external layer in CVM is a complete network, and each node in the external layer is linked with her prototype node in the internal layer. Moreover, there are no connections between nodes in the internal layer. Therefore, no internal interaction between agents is assumed in CVM.

* The work of the second author was supported by Russian Science Foundation, grant N 22-11-00051, https://rscf.ru/en/project/22-11-00051/

In paper [5] GCVM is introduced, where it is supposed that the individuals can interact in the internal layer, which is motivated by the wish of individuals to share their real opinions with their close friends.

Microscopic models directly describe how individuals' opinions evolve social individuals' perspectives, e.g., see the DeGroot model [8], the Friedkin — Johnsen (F-J) model [9], and bounded confidence models [10, 11].

In the DeGroot model, each individual updates his opinion based both on his own and his neighbors' opinions. The F-J model is an extension of the DeGroot model in which stubborn agents are considered. In the F-J model, actors can also take into account their initial prejudices in every iteration of opinion [12]. The possibility to control the agents' opinions by nonmembers of network is considered in [13, 14]. The upgraded F-J model with passive and active agents is introduced in [15]. A bounded confidence model (BCM) is a model, in which agents ignore the opinions that are very far from their own ones [6]. The BCM includes two essential models: the Deffuant — Weisbuch model (D-W) proposed in paper [11], and the Hegselman — Krause (H-K) model introduced in the work [10]. In the D-W model, two individuals a_i, a_j are randomly chosen, and they determine whether to interact according to the bounded confidence [16]. The comprehensive survey [17] examines various models in the bounded confidence opinion dynamics domain, highlighting key mechanisms leading to consensus emergence, polarization, and fragmentation within groups.

According to [18], opinion dynamics models are usually composed of a few essential elements: (i) opinion expression formats defining how to represent opinion mathematically, (ii) fusion rule determining how individuals interact with each other, and (iii) opinion dynamics environments, that is, a structure of such a social network.

In a social network, individuals neither fully accept nor completely ignore the opinions of other individuals. To a certain extent, they consider these opinions in forming their new opinions in a process defined by a fusion rule. Through a group interaction, individuals continuously update and integrate their opinions on the same issue. Eventually, there are three varieties of stabilized fusion results: consensus, polarization, fragmentation, and one unstable fusion result, that is oscillation [10].

The GCVM [5] belongs to the group of macroscopic models, we start with a network structure and use statistical-physical methods and probability theory to formulate and simulate the opinion dynamics process (i. e. in the simulations, we do not create a real network and simulate this model based on formulas).

In this paper, we examine several real networks and simulate the opinion dynamics on these networks. First, for the given internal and external structures, we create the corresponding networks. Then, we initialize the initial opinion for each individual/agent based on some parameters. Like in a macroscopic version of GCVM, we allow players to exchange their opinions with players from internal layer.

The difference between macro and micro versions of GCVM is that in a micro version model we do not need to adjust the simulation program according to a different network structure. As long as a network structure is given, the program will automatically produce simulations. Therefore, it will be convenient to use this program to simulate a real network structure. But in a macro version, for different network structures, we should adjust the corresponding state transition formulae.

The primary conclusion drawn from this research indicates that a cyclic external structure invariably increases consensus time and positively impact a winning rate. Furthermore, it has been observed that when a circle in the external layer is extended to a complete graph, it significantly impacts the consensus time due to the reduction of the average shortest path between every pair of nodes in the external layer (see formula (2)). While each parameter individually influences consensus time, none of them impact the winning rate.

The rest of this paper is organized as follows. Section 2 introduces a model. In Section 3 we present the experiments and results. The conclusions are given in Section 4.

2. Model. A two-layer network with N individuals/agents is defined by

• N: number of individuals/agents in the network;

• $a_i = (a_i^E, a_i^I)$: individual/agent i, where $i = 1, \ldots, N$, a_i^E (a_i^I) is a representation of agent i in the external (internal) layer (i.e. a set of individuals/agents is the same for both layers);

• $G_E(\mathcal{V}_E, \mathcal{E}_E)$: predefined external network, where $\mathcal{V}_E = \{a_i^E\}$, $i = 1, \ldots, N$, represents a set of individuals and $\mathcal{E}_E$ — a set of social relations between individuals in the external layer;

• $G_I(\mathcal{V}_I, \mathcal{E}_I)$: predefined internal network, where $\mathcal{V}_I = \{a_i^I\}$, $i = 1, \ldots, N$, represents a set of individuals and $\mathcal{E}_I$ — a set of social relations between individuals in the internal layer;

• $\mathcal{E}_C = \{(a_i^E, a_i^I) | i = 1, \ldots, N\}$: set of edges connecting individuals in external and internal layers.

We define a two-layer network with N individuals/agents as

$$G(\mathcal{V}, \mathcal{E}), \tag{1}$$

where $\mathcal{V} = \mathcal{V}_E \cup \mathcal{V}_I$, $|\mathcal{V}_E| = |\mathcal{V}_I| = N$, and $\mathcal{E} = \mathcal{E}_E \cup \mathcal{E}_I \cup \mathcal{E}_C$. This definition is independent of a specific network structure, i.e. external/internal networks can be different.

2.1. The general concealed voter model (macro version). Zhao and Parilina [5] proposed GCVM based on CVM introduced in [3]. These papers use simulations to represent opinion transmission processes in two-layer networks. In the following section, we introduce GCVM in a micro version.

2.2. The general concealed voter model (micro version). In the GCVM, we use R, B (r, b) to represent individuals' external (internal) red and blue opinions respectively. There is a list of notations:

• $S = \{Rr, Rb, Br, Bb\}$: set of all possible states of an individual;

• $\omega(a_i, t) \in S$: opinion of individual a_i at time t, where $i = 1, \ldots, N$, and $t = 0, 1, \ldots$;

• ρ_{r_e}: ratio of individuals having red opinion in external layer;

• ρ_{r_i}: ratio of individuals having red opinion in internal layer;

• ρ_r: ratio of individuals having red opinion in both internal and external layers;

• r_e: number of individuals having red opinion in external layer;

• r_i: number of individuals having red opinion in internal layer;

• r: number of individuals having red opinion in both internal and external layers;

• π_{c_e}: external copy rate, that is a probability of an individual to copy opinion of his/her external neighbor;

• π_{c_i}: internal copy rate, that is a probability of an individual to copy opinion of his/her internal neighbor;

• π_e: externalization rate, that is a probability of hypocrisy* choosing to publicly express his/her internal opinion;

* By hypocrisy we mean a node having different opinions in external and internal layers, i.e., the nodes in states Rb and Br.

• π_i: internalization rate, that is a probability of hypocrisy accepting his/her external opinion.

In Sections 2.2.1–2.2.3, we describe GCVM of opinion dynamics in a two-layer network. The description is organized so that to understand how numerical simulations presented in Section 3 are done. In the following section, we introduce GCVM in a micro version.

2.2.1. Two-layer network structure initialization. We start by setting two networks G_E and G_I (we read these networks from the file, and an example of such a file representing external cycle and internal star structures is shown in Listing. Then we add the edges between external and internal representations of individuals).

Listing. **Example of graph file with external cycle and internal star structure**

```
>>>external
E0 E1
E1 E2
E2 E3
E3 E0
<<<external end
>>>internal
I00 I1
I00 I2
I00 I3
<<<internal end
```

This results in a two-layer network G we store as an adjacency list.

2.2.2. Initialization of individuals' initial states. Denote a number of individuals in state $s \in S$ by $\#s$. We have the following relations:

$$\begin{aligned}
N &= \#Rr + \#Rb + \#Br + \#Bb,\\
r_e &= \#Rr + \#Rb,\\
r_i &= \#Rr + \#Br,\\
r &= \#Rr,\\
\#Bb &= N - r_e - r_i + r,\\
\#Rb &= r_e - r,\\
\#Br &= r_i - r.
\end{aligned}$$

Assuming a uniform distribution for each agent to belong to any state $s \in S$ at the initial time, we adopt the following rule of setting the initial state $\omega(a_i, 0)$ for any agent a_i at time $t = 0$:

$$\omega(a_i, 0) := f(x) = \begin{cases} Rr, & 0 \leqslant x < \rho_r,\\ Rb, & \rho_r \leqslant x < \rho_{r_e},\\ Br, & \rho_{r_e} \leqslant x < \rho_{r_e} + \rho_{r_i} - \rho_r,\\ Bb, & \rho_{r_e} + \rho_{r_i} - \rho_r \leqslant x \leqslant 1, \end{cases}$$

where $x \sim U(0, 1)$.

2.2.3. Opinion transmission process. We can divide individuals into hypocrites and nonhypocrites based on the consistency of their external and internal opinions. Hypocrites are individuals who have different opinions in the internal and external layers, while non-hypocrites have the same opinions in both layers.

We focus on two measurements to analyze GCVM:

• *consensus time:* T_{cons} is consensus time in (G)CVM, that is, the time required for all individuals to form the same opinion in internal and external layers (i.e., $\rho_{r_e} = \rho_{r_i} = \rho_r = 0$ or 1 for T_{cons});

• *winning rate:* ρ is a winning rate of red opinion in a series of simulations. For the opinion, to win means that there is no other opinion that agents have in the whole network (i.e. in a series of simulations, the number of simulations, in which red opinion wins blue opinion divided by the number of simulations).

Before presenting an algorithm of GCVM, we briefly define the actions available for a randomly chosen individual a_i:

• *picking up a_i's neighbor*: randomly choose a neighbor among all a_i's neighbors. Let it be individual a_j (this is a prerequisite action for external/internal copying);

• *external copying:* a_i copies a_j's external opinion with probability π_{c_e};

• *internal copying:* a_i copies a_j's internal opinion with probability π_{c_i};

• *externalization:* a_i expresses his/her internal opinion with probability π_e (this action is available only for hypocrite);

• *internalization:* a_i accepts his/her external opinion with probability π_i (this action is available only for hypocrite).

Externalization and internalization are meaningless for non-hypocrites, so they have only two possible actions (external and internal copies).

***Algorithm of GCVM*:**

Step 1. Initialize $t = 0$.

Step 2. Choose an individual a_i, uniformly random from N individuals in two-layer network G.

Step 3. Check all valid actions of individual a_i (depending on his/her state) and randomly choose one of the valid actions with equal probabilities:

I) a_i is a hypocrite, then he/she has four possible actions: (i) external copying, (ii) internal copying, (iii) externalization, and (iv) internalization. Any action is chosen with a probability of 0.25;

II) a_i is a nonhypocrite, then he/she can perform only external or internal copying. Any action is chosen with a probability of 0.5.

Step 4. Generate random number $x \sim U(0, 1)$. Perform the action chosen in Step 3:

a) if external copying is chosen in Step 3 and $x < \pi_{c_e}$, then a_i copies a_j's external opinion;

b) if internal copying is chosen in Step 3 and $x < \pi_{c_i}$, then a_i copies a_j's internal opinion;

c) if externalization is chosen in Step 3 and $x < \pi_e$, then a_i expresses his/her internal opinion;

d) if internalization is chosen in Step 3 and $x < \pi_i$, then a_i accepts his/her external opinion.

Step 5. Increase t by 1. If consensus is reached*, stop iteration. Otherwise, go back to Step 2.

3. Experiments and results.

3.1. General description. The experiment focuses on observing the effect of an external network structure on a winning rate of opinion and consensus time.

* The algorithm will be stopped when all individuals in both layers hold the same opinion, i. e. consensus is reached.

We consider two types of external structures: cycle and complete, and seven internal structures: cycle, complete, star, two-star (odd and even cases), and two-clique (odd and even cases). This gives us 14 different combinations of external-internal structures, i. e. 14 two-layer networks.

Figure 1 is an example of a two-layer network with an external cycle and internal two-clique layers. For other structures, we need to modify the corresponding layer accordingly, which we do not present here to save the space.

a

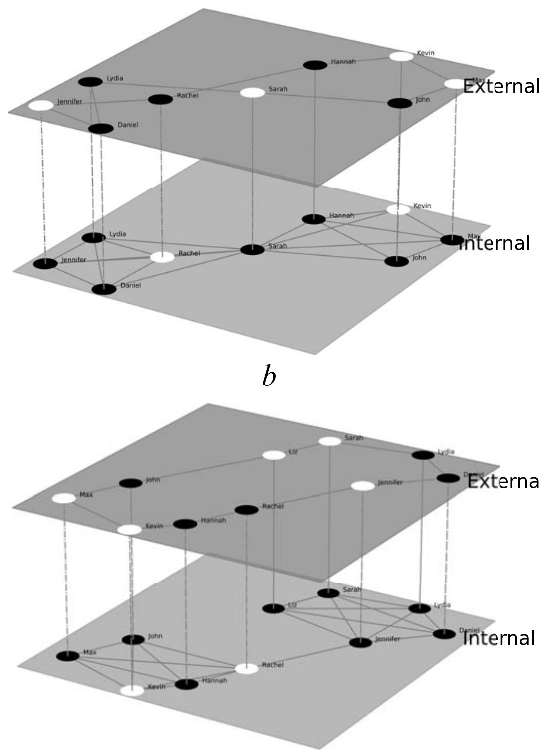


Figure 1. Representation of external cycle with internal two-clique network
Black represents red opinion, and white represents blue opinion.
a — odd number of nodes; *b* — even number of nodes.

For our experiments, we fix the following parameters: $\rho_{r_e} = 0.75$, $\rho_{r_i} = 0.25$, $\rho_r = 0.2$, $\pi_{c_i} = \pi_{c_e} = 1$, $\pi_e = 0.01$, $\pi_i = 0.5$, $N = 100$. Then we observe the effect of external structure on the winning rate of opinion and consensus time (number of iterations) for the given internal structure. We conduct 100 simulations for each model and obtain the following statistical results. The models presented in Figure 2, *a* and *b* below are named as "external layer — internal layer — # of individuals".

As shown* in Figure 2, *a*, we conclude that cyclic external structures prolong consensus time in comparison with complete external structure (the similar result is obtained in [5]). This conclusion is true for all internal structures we examine in the experiment.

* The number in the model name represents the number of individuals. e.g. 50 represents an even case, and 51 is an odd case.

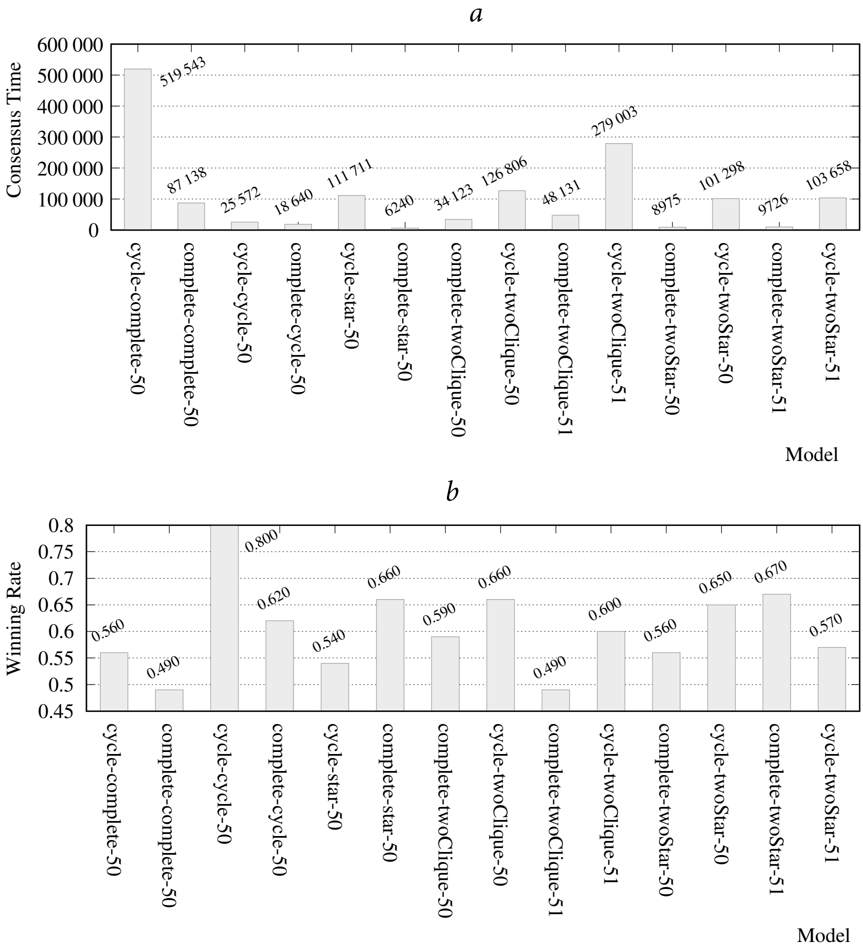


Figure 2. Observed consensus time (a) and winning rate (b) for models with different two-layer network structures

Additionally, we observe that the internal structure also has an impact on consensus time. For instance, consensus time for a two-star internal structure is less than that of a two-clique internal structure.

However, the observation results for the winning rate are quite different than in our previous work [5], as shown in Figure 2, *b*. We can notice that except internal structures "twoStar-51" and "star-50", a cyclic external layer decreases the winning rate. For all other models, a cyclic external structure has a positive impact on the winning rate. The possible reason is in specification of a microscopic model, i.e. actions that an individual/agent can take are related only to his current state*. In a macroscopic version of GCVM, the probability of each possible action is related to the overall state of the system represented by a triple $(\rho_{r_e}, \rho_{r_i}, \rho_r)$ (see [5]).

* For nonhypocrites, an individual has two possible actions, and for hypocrites, an individual has four possible actions. The probability for each possible action of an individual/agent at the present moment is fixed.

3.2. Main results and observations. Based on the findings from Section 3.1 indicating that a cyclic structure has a positive impact on consensus time, we have formulated the following research questions. A series of experiments were designed and conducted in order to address these research questions:

1. How does an external structure influence KPIs, i.e. are there any features of a network that significantly affect KPIs? To address these question we do the following:

a) extend a cyclic structure to a complete one in different ways*,

b) reduce a cyclic structure to a line, and observe the KPIs,

c) conduct statistical tests to determine if there are significant differences in the distribution of KPIs for different ways of constructing a complete graph (see Item a),

d) check the correlation between network features and KPIs.

2. How do externalization and internalization rates influence KPIs?

a) avary π_e from 0.1 to 1.0 with a step of 0.1,

b) vary π_i from 0.3 to 1.0 with a step of 0.1 (note: when $\pi_i = 0.1$ or $\pi_i = 0.2$, consensus cannot be reached in some models).

3. How does a copying rate influence KPIs?

a) vary π_c from 0.1 to 1.0 for both layers with a step of 0.1.

4. How does a combination of parameters influence KPIs (i.e. which combinations maximize or minimize consensus time and winning rate)?

a) use a combination of parameters $(\pi_i, \pi_e, \pi_{c_i}, \pi_{c_e})$, where we vary $\pi_i \in [0.3, 1]$ and $\pi_e, \pi_{c_i}, \pi_{c_e} \in [0.1, 1]$ with a step of 0.1.

We start by extending a cyclic structure in the following three different ways:

• *normal*: find a set of edges presented in a complete graph but not in a cyclic graph, and add them sequentially to a cyclic graph until finally obtain a complete graph;

• *random*: find a set of edges presented in a complete graph but not in a cyclic graph, and add them randomly to a cyclic graph;

• *shortest*: find a set of edges presented in a complete graph but not in a cyclic graph, and add the edges from the list that minimizes d, where d is the average of the shortest paths among all pairs of nodes in external layer. Here d is calculated as follows:

$$d = \sum_{s,t \in \mathcal{V}_E} \frac{d(s,t)}{n_E(n_E - 1)}, \tag{2}$$

where $d(s,t)$ is the length of the shortest path between s and t, $\mathcal{V}_E$ is a set of nodes in external layer, $n = |\mathcal{V}_E|$ is the number of nodes in external layer.

Obviously, for an undirected graph with N nodes, a cyclic structure has only N edges, and a complete structure has $N \times (N-1)/2$ edges. Therefore, a cyclic graph can become a complete graph by adding $N \times (N-3)/2$ edges. If we remove one edge from a cyclic strucutre, it will degenerate to a line.

In our experiments, we extended a network structure from ‘cycle-complete-50’ to ‘complete-complete-50’, defined ‘cycle+δ-complete-50’ as an intermediate network structure, where $\delta \in \mathbb{Z}$ means the number of edges which we have added to a cyclic structure by an iteration. Here $\delta \in [-1, 1175]$, and the value $\delta = -1$ corresponds to the case when we deleted an edge from a cycle degenerating this cycle into a line. When $\delta = 1175$, the cycle becomes a complete graph, i.e. the maximal number of edges that can be added into a cycle with 50 nodes is equal to 1175. In our experiments we examined the dynamics of consensus time and d when δ is increasing.

* A cycle graph can be transformed into a complete graph by adding several nonexistent edges, or degenerate into a line by deleting an existing edge.

The results of simulations are shown in Figure 3. In particular, Figure 3, a shows how d (δ) influences winning rate for different ways of extending a cyclic structure to a complete one. The right figure is an increase of the first 150 points from the left graph, we did the same in Figures 3, b and 4. Looking at Figure 3, a, we can notice that winning rate is almost white noise with changes of δ*.

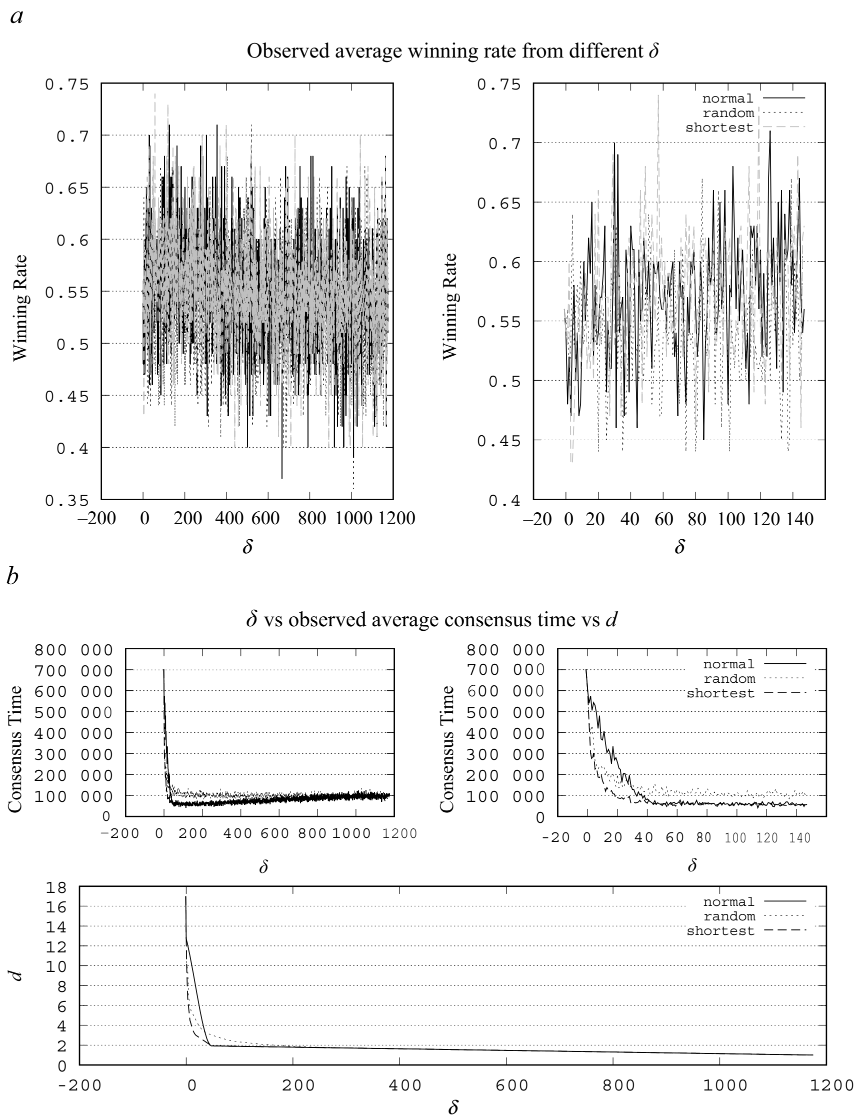
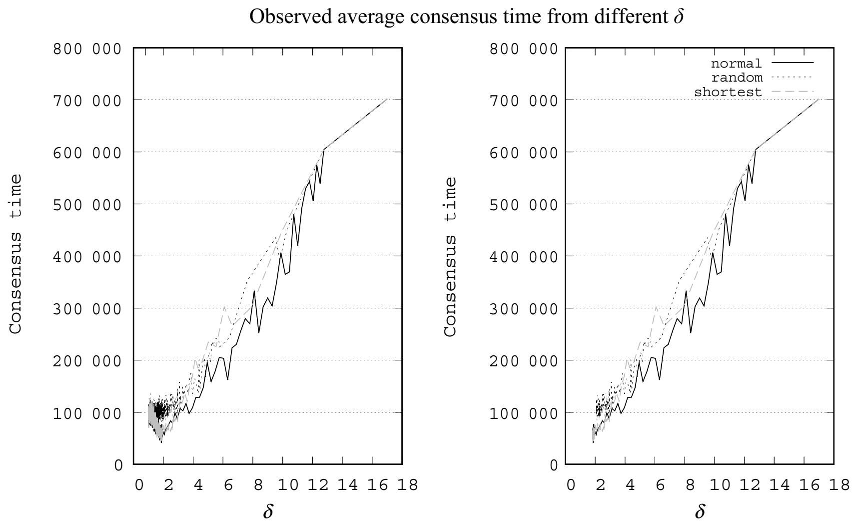


Figure 3. Winning rate (a) and consensus time (b) for the models with different extension ways

Figure 3, b shows a trend of consensus time when δ is changing, and it is easy to recognize that the black solid and gray dashed lines have almost the same trend. At

* The winning rate is white noise for "random" with lag 1–10, for "shortest" with lag 1–3, but the winning rate for "normal" extending way is not white noise which is confirmed by Ljung—Box test.

a



b

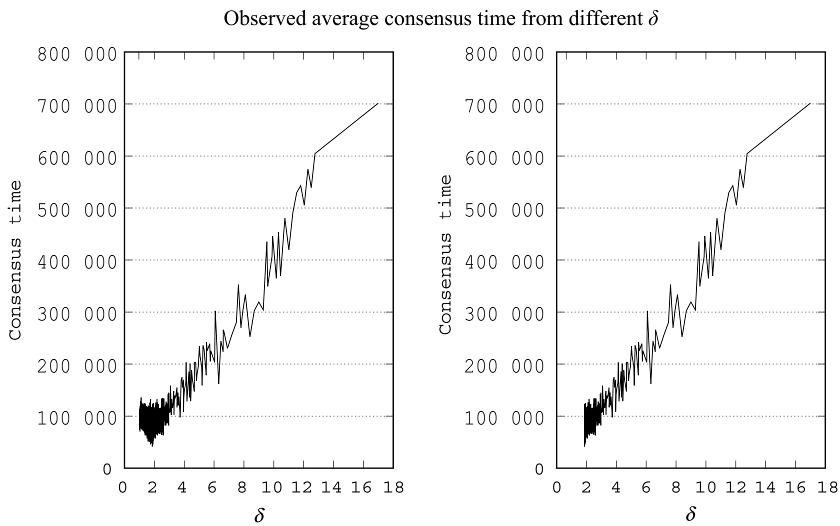


Figure 4. Consensus time vs d
a — for different extending ways; b — aggregated data.

the beginning, with an increase of the number of added edges, consensus time significantly decreases. After reaching a certain level, the increase in the number of edges has no significant impact on consensus time. First points in Figure 3, *b* correspond to a network structure “cycle+-1-complete-50” having an external line structure. It is obvious that $d_{\text{line}} > d_{\text{cycle}} > d_{\text{complete}}$. Therefore, we formulate Hypothesis 1.

Hypothesis 1. *There is a significant correlation between consensus time and* d.

In order to verify Hypothesis 1, we should find d for each graph in Figure 3, *b*. After calculating d for each graph, we construct the third subfigure in Figure 3, *b*. Obviously, the first and third subfigures have a similar trend.

We represent the relation between d and consensus time in Figure 4. The difference between Figure 4, a and b is that in Figure 4, a we group the data by extending way first, and then draw the trend, while in Figure 4, b, we do not specify the way of extending the graph, but only make an analysis based on different d and consensus time.

We can observe an approximately linear relationship in Figure 4. We can further use statistical methods for correlation analysis [19, 20]. The results of examining the Pearson correlation coefficient (PCC) [21] are shown in Figure 5. We make the following conclusions:

• the correlation between d and consensus time is significantly strong, and PCC is 0.78;

• for other pairs of KPIs, the correlation is not significant, and the absolute values of PCC are less than 0.15.

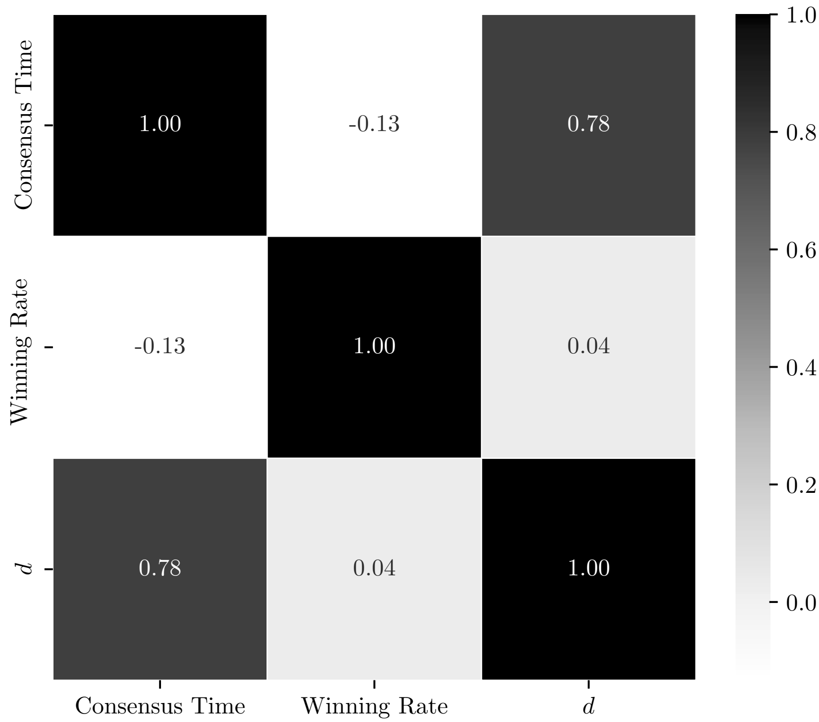


Figure 5. Pearson correlation coefficients

Hypothesis 2. *There are significant differences in KPI distributions for different ways of constructing a complete graph.*

We are interested in how a different extending way influences the distribution of KPIs. We show empirical distributions of KPIs in Figure 6. We should notice that their distributions are significantly different for different ways of extending the graph from a circle to a complete one. But for 'normal' and 'shortest' extending ways, KPIs distributions are very similar. We use the Kolmogorov — Smirnov test for further analysis [22, 23]. The results are shown in Table 1. From Table 1 we can see that p-values for all KPIs when we compare normal and shortest extending ways are larger than 0.05. We make the following conclusion: we should accept the null hypothesis that the distributions of KPIs for normal and shortest extending ways are identical.

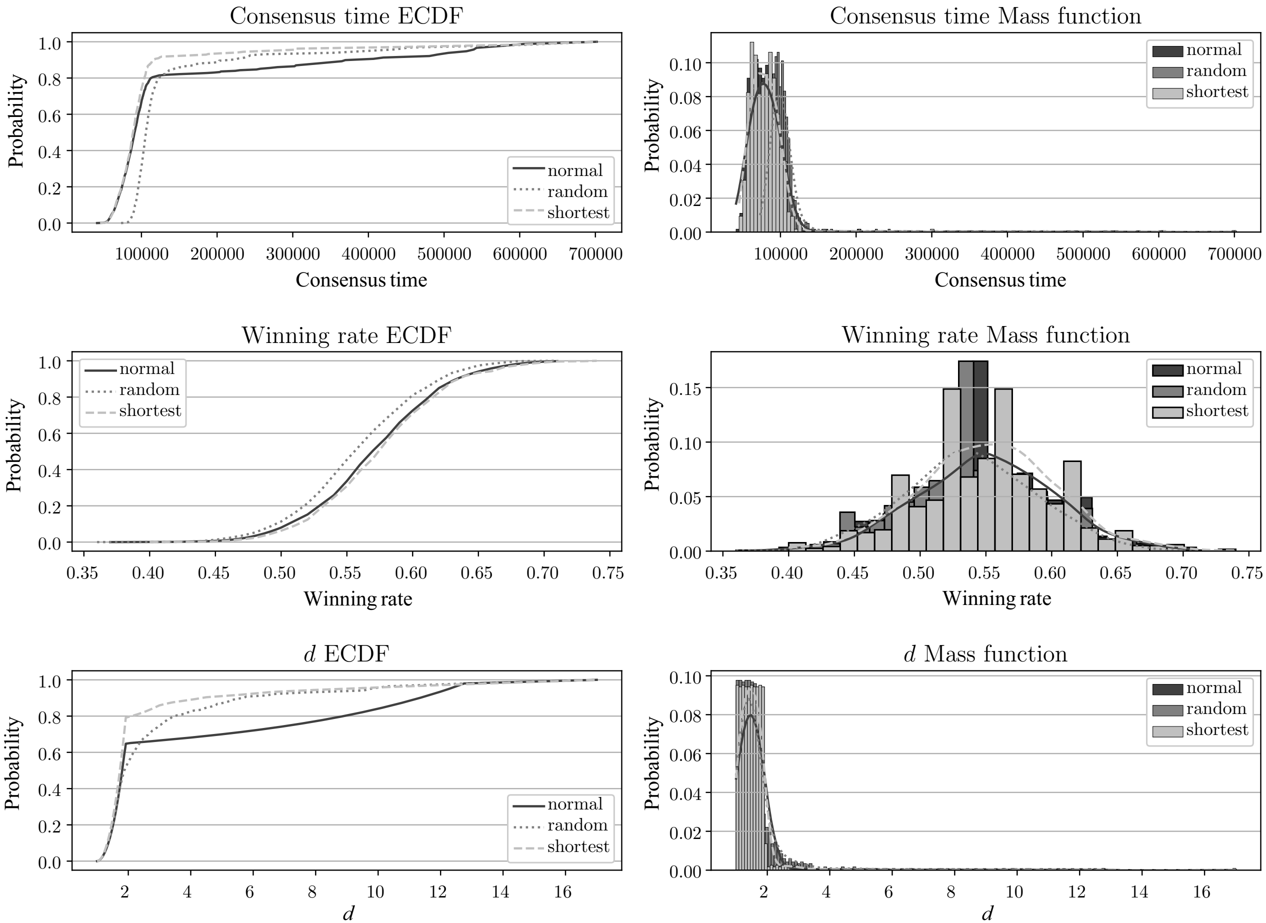


Figure 6. Frequency mass function and empirical cumulative distribution function (ECDF) for different KPIs

Table 1. **Results of Kolmogorov — Smirnov tests**

Pair	Statistics	p-values
Consensus time		
normal vs random	0.557	0.0
normal vs shortest	0.031	0.641
random vs shortest	0.574	0.0
Winning rate		
normal vs random	0.114	0.0
normal vs shortest	0.02	0.967
random vs shortest	0.121	0.0
d		
normal vs random	0.121	0.0
normal vs shortest	0.02	0.978
random vs shortest	0.12	0.0

For now we find out that some ways of extending a circle to a complete graph have an impact on KPI distributions. But how significant is this impact? How are mean and variance affected? Therefore, we formulate next hypothesis.

Hypothesis 3. *Means and variances of KPIs are the same for different ways of extending a circle to a complete graph.*

We use some statistical tests to verify equity of variances and equity of means. Before doing this, we first run normality tests [24, 25] since some statistical tests are parametric, i.e. they assume normality of the data. The results of normality tests are shown in Table 2, where we can see that the p-value for all KPIs are smaller than 0.05. Then we should reject the null hypothesis that any KPI is normally distributed.

Table 2. **Results of normality tests**

KPIs		Statistics	p-values
Normality testing overall			
Consensus time		0.5	0.0
Winning rate		0.996	0.0
d		0.396	0.0
Mode	KPIs	Statistics	p-values
Normality testing grouped via mode			
normal	Consensus time	0.379	0.0
normal	Winning rate	0.996	0.003
normal	d	0.32	0.0
random	Consensus time	0.347	0.0
random	Winning rate	0.996	0.003
random	d	0.469	0.0
shortest	Consensus time	0.442	0.0
shortest	Winning rate	0.995	0.001
shortest	d	0.369	0.0

In Tables 3 and 4, we have two group of results, 'ev/em test for all' corresponds to whether the variances/means of three extension ways are all equal. The 'pairwise ev/em test' corresponds to two-sample equity test of variances/means. As none of KPIs is normally distributed, we use the Levene test for variance equity [26–29].

We make the following conclusions from Table 3:

• the variance of winning rate is the same for all extension ways (all p-values in the Levene tests are greater than 0.05);

• we reject the null hypothesis that the variances of the consensus time are equal for all extension ways (all p-values in the Levene tests are less than 0.05);

• variances of d are equal for normal and random extension ways.

Table 3. **Results of variance equity tests**

Test	Description	KPIs	Statistics	p-values
EV test for all				
Fligner test	Distribution free	Consensus time	195.249	0.0
	when populations	Winning rate	1.206	0.547
	are identical	d	23.341	0.0
Levene test	More robust	Consensus time	12.533	0.0
	for significantly	Winning rate	1.074	0.342
	non-normal population	d	5.441	0.004
Bartlett test	More depends	Consensus time	320.877	0.0
	on normal	Winning rate	3.584	0.167
	population	d	351.795	0.0
Pairwise EV test				
	normal vs random	Consensus time	20.967	0.0
	normal vs shortest	Consensus time	6.579	0.01
	random vs shortest	Consensus time	7.651	0.006
Levene test	normal vs random	Winning rate	1.572	0.21
for pairs	normal vs shortest	Winning rate	0.003	0.959
	random vs shortest	Winning rate	1.71	0.191
	normal vs random	d	0.187	0.665
	normal vs shortest	d	8.533	0.004
	random vs shortest	d	10.099	0.002

Since none of KPIs is normally distributed, and not all KPIs are homoscedastic, to verify equity of means we use the Kruskal test [30–33]. The results of the tests are given in Table 4 and we conclude the following:

• the means of d are equal for all extension ways, i.e. a way of extending a circle to a complete graph does not affect the mean of d (all p-values in the Kruskal tests are greater than 0.05). For consensus time and winning rate, the means are not all equal, i.e. they differ by extension ways;

• we accept the null hypothesis that the means of consensus time (and winning rate) are equal for normal and shortest extension ways (all p-values in KruskalResult are greater than 0.05).

Figures 7 and 8 show how parameters (π_{c_e} and π_{c_i} in Figure 7, and π_e and π_i in Figure 8) influence winning rate. We can see that winning rate fluctuates within a certain range, but not too much. Therefore, we temporarily think that an impact of parameters on winning rate is limited.

Figures 9 and 10 show how consensus time varies with a change of parameters (π_{c_e} and π_{c_i} in Figure 9, and π_e and π_i in Figure 10). We make these interesting observations:

• an increase of external copying rate π_{c_e} has a negative effect on consensus time. The interpretation is as follows: when an individual in a society is more inclined to listen to the opinions of his/her external neighbors, it is helpful to reach consensus;

• with an increase of internal copying rate π_{c_i}, consensus time increases;

• with an increase of externalization rate π_e, consensus time first increases until it reaches the maximal value, and then decreases. The interpretation of this is as follows: expressing your true opinion to a certain extent is not effective to reach consensus within the whole system, but beyond this threshold, along with an increase of desire to express your opinion, for the system, it is easier to reach consensus;

• with an increase of internalization rate π_i, consensus time decreases. This can be interpreted as follows: when people are more willing to accept their own external opinion, it will accelerate consensus of the whole system.

In Tables 5 and 6, we show minimal/maximal consensus time and winning rate re-

Table 4. **Results of mean equity tests**

Test	Description	KPIs	Statistics	p-value
EM test for all				
f_oneway test	Independent sample; each sample is from a normally distributed population; homoscedasticity	Consensus time	99.143	0.0
		Winning rate	23.205	0.0
		d	4.795	0.008
Kruskal test	Sample size should > 5	Consensus time	1100.711	0.0
		Winning rate	46.919	0.0
		d	1.596	0.45
Alexander Govern test	Independent sample; each sample is from a normally distributed population; heteroscedasticity	Consensus time	261.282	0.0
		Winning rate	47.419	0.0
		d	12.243	0.002
Pairwise EM test				
Kruskal test for pairs	normal vs random	Consensus time	795.081	0.0
	normal vs shortest	Consensus time	0.775	0.379
	random vs shortest	Consensus time	854.985	0.0
	normal vs random	Winning rate	32.215	0.0
	normal vs shortest	Winning rate	0.229	0.632
	random vs shortest	Winning rate	37.928	0.0
	normal vs random	d	1.092	0.296
	normal vs shortest	d	0.001	0.972
	random vs shortest	d	1.301	0.254

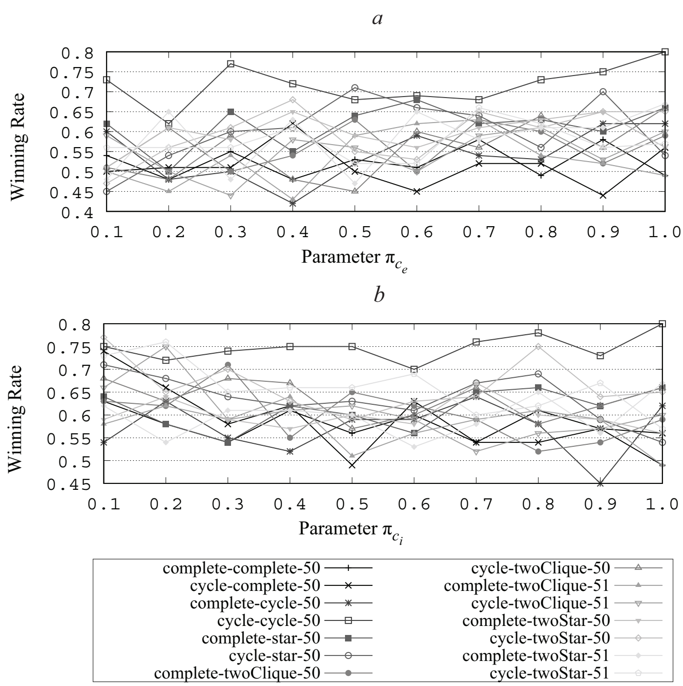


Figure 7. Winning rate for different copying rates
a — varying external copying rate π_{c_e}; b — varying internal copying rate π_{c_i}.

spectively for different combinations of parameters. If we compare the left and right parts in Tables 5 and 6, there is a large difference between maximum and minimum values.

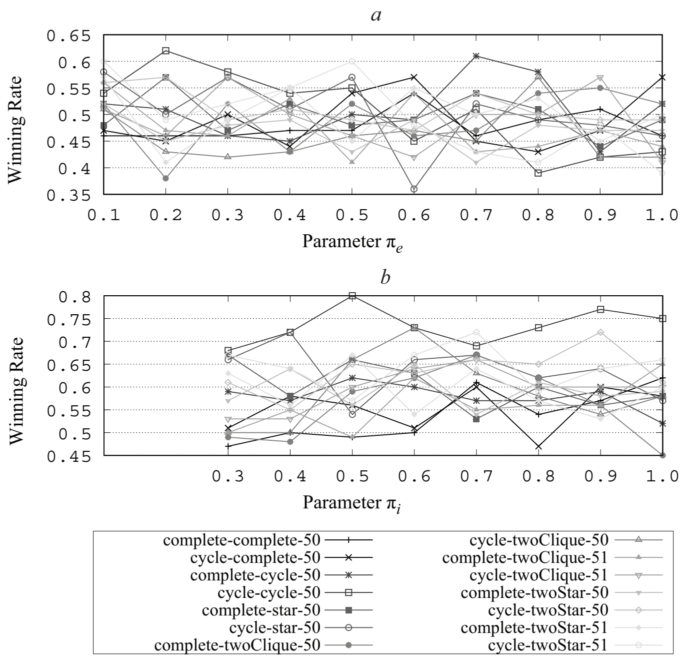


Figure 8. Winning rate for different externalization and internalization rates
a — varying externalization rate π_e; b — varying internalization rate π_i.

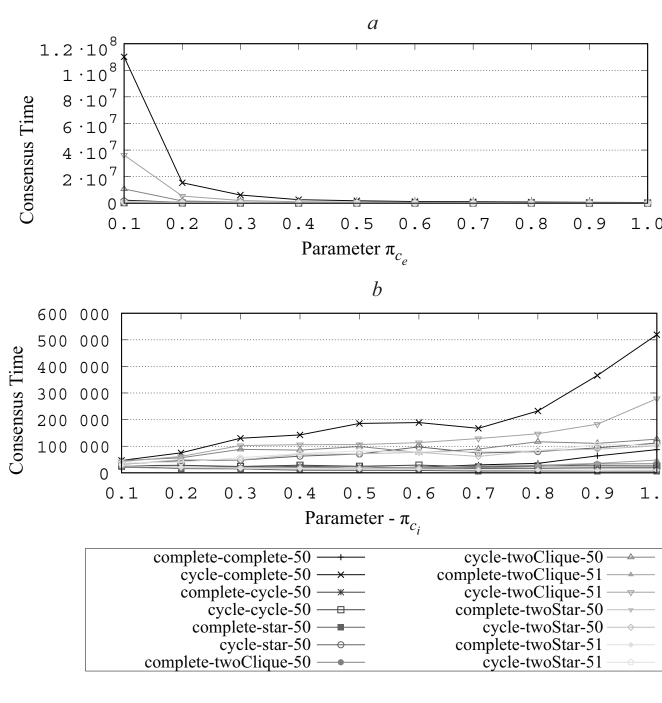


Figure 9. Consensus time for different copying rates
a — varying external copying rate π_{c_e}; b — varying internal copying rate π_{c_i}.

Table 5. **Observed minimum and maximum consensus time**

External	Internal	N	Max	π_{c_e}	π_{c_i}	π_e	π_i	Min	π_{c_e}	π_{c_i}	π_e	π_i
complete	complete	50	1 185 261 644	0.1	1.0	0.1	0.3	28767	0.8	0.1	0.1	1.0
cycle	complete	50	1 144 248 878	0.6	1.0	0.1	0.3	29387	0.9	0.1	0.3	1.0
complete	cycle	50	26 467 498	1.0	1.0	0.4	0.3	22850	1.0	0.1	0.1	1.0
cycle	cycle	50	474 773	0.1	0.1	0.6	0.9	26842	1.0	0.9	0.1	1.0
complete	star	50	1602 538	1.0	0.1	0.5	0.3	2018	0.1	1.0	1.0	0.3
cycle	star	50	275 898	1.0	1.0	0.1	0.3	2021	0.1	1.0	1.0	0.3
complete	twoClique	50	492 816 970	0.2	1.0	0.1	0.3	28185	0.1	0.7	1.0	0.3
cycle	twoClique	50	151 026 217	0.7	1.0	0.1	0.3	20923	0.1	1.0	1.0	0.3
complete	twoClique	51	708 728 030	0.1	1.0	0.1	0.3	30628	0.9	0.1	0.1	1.0
cycle	twoClique	51	449 651 392	0.6	1.0	0.1	0.3	26250	0.1	1.0	1.0	0.3
complete	twoStar	50	2027 297	1.0	0.1	0.5	0.3	3795	0.1	0.9	1.0	0.3
cycle	twoStar	50	317 571	1.0	0.9	0.1	0.3	3743	0.2	1.0	1.0	0.4
complete	twoStar	51	2 347 437	1.0	0.1	0.5	0.3	4201	0.1	0.9	1.0	0.3
cycle	twoStar	51	349 638	0.7	1.0	0.1	0.3	4527	0.2	1.0	1.0	0.3

Table 6. **Observed minimum and maximum winning rate**

External	Internal	N	Max	π_{c_e}	π_{c_i}	π_e	π_i	Min	π_{c_e}	π_{c_i}	π_e	π_i
complete	complete	50	0.7	0.1	0.1	0.2	1.0	0.31	0.2	0.6	1.0	0.5
cycle	complete	50	0.81	0.9	0.1	0.2	0.8	0.33	0.4	0.1	0.8	0.3
complete	cycle	50	0.71	0.2	0.1	0.4	1.0	0.29	0.7	1.0	0.6	0.3
cycle	cycle	50	0.82	0.1	0.6	0.1	0.6	0.25	0.2	0.9	1.0	0.6
complete	star	50	0.7	0.1	0.4	0.2	1.0	0.24	0.3	0.4	1.0	0.4
cycle	star	50	0.82	0.2	0.1	0.1	0.7	0.23	0.1	0.2	1.0	0.4
complete	twoClique	50	0.68	0.5	0.7	0.4	0.8	0.3	0.7	0.3	0.9	0.9
cycle	twoClique	50	0.76	0.9	0.1	0.1	1.0	0.28	0.2	0.8	0.6	0.5
complete	twoClique	51	0.73	0.1	0.1	0.1	0.9	0.32	0.4	0.8	0.7	0.4
cycle	twoClique	51	0.77	0.7	0.2	0.1	0.9	0.31	0.1	0.3	1.0	0.5
complete	twoStar	50	0.72	0.1	0.4	0.2	1.0	0.26	0.1	0.8	0.8	0.4
cycle	twoStar	50	0.77	0.8	0.2	0.2	0.9	0.25	0.3	0.3	1.0	0.3
complete	twoStar	51	0.72	0.1	0.1	0.1	0.7	0.27	0.1	0.8	0.9	0.3
cycle	twoStar	51	0.8	0.8	0.2	0.1	1.0	0.26	0.1	1.0	0.9	0.4

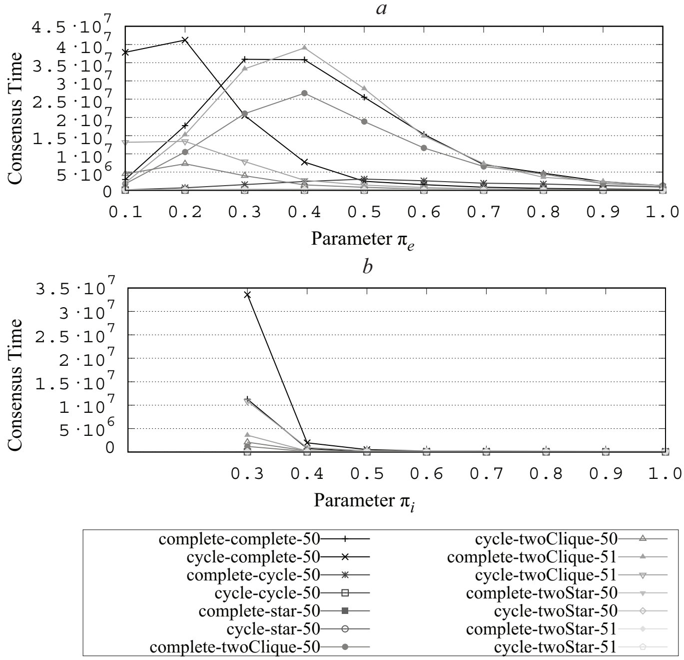


Figure 10. Consensus time for different externalization and internalization rates
a — varying externalization rate π_e; b — varying internalization rate π_i.

We performed a clustering procedure using K-means method with input variables being consensus time and winning rate separately [34, 35]. The resulting cluster labels are then added to the original data. Based on these cluster labels, we observed the distribution of the corresponding four parameters (π_{c_e}, π_{c_i}, π_e, and π_i).

Ideally, we prefer having two clusters since it allows us to determine which parameter combinations result in respectively large or small KPIs. In practice, we specify a range for the number of clusters k, from 1 to 20, and calculate the silhouette score [36]. In some sense an optimal value of k is the one maximizing the silhouette score.

Since the distributions of winning rates are very close to a normal distribution, based on clustering results of winning rate, we cannot observe any significant differences in descriptive statistics of clusters (see Table 7). Therefore, in our future analysis we only focus on performing a clustering analysis of consensus time.

Table 8 shows the number of elements in the clusters of consensus time, where we selected to have two clusters since it maximizes the silhouette score.

In Figure 11, a, the distributions of consensus time are almost consistent for different multi-layer models. At the same time, in Figure 11, b, we can clearly observe noticeable differences. This interesting result provides us with a valuable insight that the diverse parameter distributions can significantly prolong consensus time.

In the top right corner of Figure 11, b, we observe that when the internal structure is complete or twoClique, parameter π_{c_i} differs significantly consensus time for these models in comparison with other models. In the lower left corner, we can see when the internal structure is complete or twoClique, a value of π_e in cluster 1 is always equal to 0.1. For the figures in the top left and lower right corners, we can get similar conclusions by comparison.

Table 7. **Descriptive statistics of clusters within winning rate (complete-complete-50)**

Parameters	Cluster	Count	Mean	Std	Min	25 %	50 %	75 %	Max
π_{c_e}	0	169	0.484	0.282	0.1	0.3	0.4	0.7	1.0
	1	238	0.540	0.287	0.1	0.3	0.5	0.8	1.0
	2	186	0.573	0.274	0.1	0.3	0.6	0.8	1.0
	3	291	0.543	0.289	0.1	0.3	0.5	0.8	1.0
	$\vdots$								
π_{c_i}	0	169	0.487	0.306	0.1	0.2	0.4	0.7	1.0
	1	238	0.591	0.293	0.1	0.3	0.6	0.9	1.0
	2	186	0.533	0.272	0.1	0.3	0.5	0.8	1.0
	3	291	0.547	0.290	0.1	0.3	0.5	0.8	1.0
	$\vdots$								
π_e	0	169	0.509	0.290	0.1	0.3	0.5	0.7	1.0
	1	238	0.581	0.304	0.1	0.3	0.6	0.9	1.0
	2	186	0.525	0.292	0.1	0.3	0.5	0.8	1.0
	3	291	0.587	0.282	0.1	0.4	0.6	0.8	1.0
	$\vdots$								
π_i	0	169	0.698	0.233	0.3	0.5	0.7	0.9	1.0
	1	238	0.632	0.230	0.3	0.4	0.6	0.8	1.0
	2	186	0.631	0.237	0.3	0.4	0.6	0.8	1.0
	3	291	0.622	0.230	0.3	0.4	0.6	0.8	1.0
	$\vdots$								

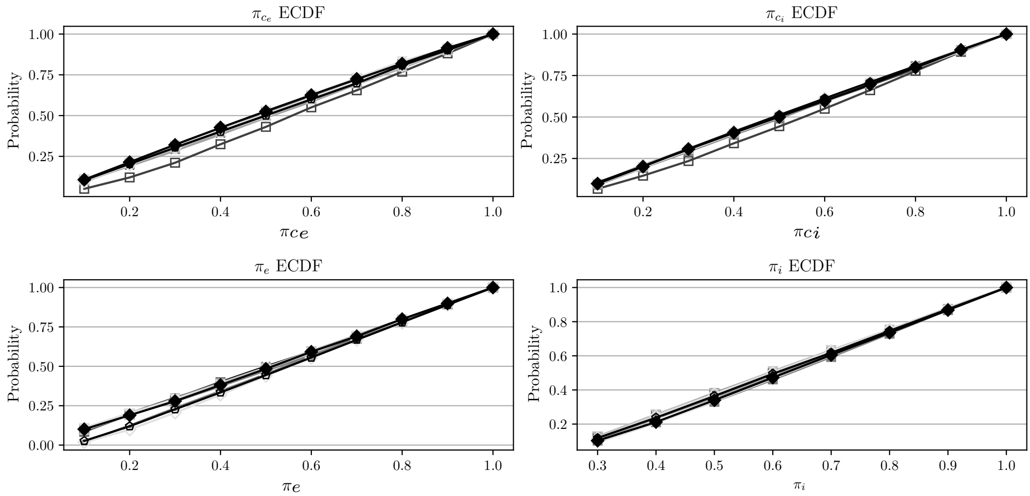
Table 8. **Consensus time cluster sizes for each model**

Model		Cluster 0	Cluster 1
0	complete-complete-50	3267	20
1	cycle-complete-50	3241	19
2	complete-cycle-50	3319	105
3	cycle-cycle-50	3027	400
4	complete-star-50	3219	205
5	cycle-star-50	3124	312
6	complete-twoClique-50	3239	24
7	cycle-twoClique-50	3348	11
8	complete-twoClique-51	3904	14
9	cycle-twoClique-51	3831	17
10	complete-twoStar-50	3171	197
11	cycle-twoStar-50	3100	372
12	complete-twoStar-51	3598	225
13	cycle-twoStar-51	3634	362

4. Conclusions. The paper introduces a novel approach to simulate GCVM by creating a real network instead of using statistical-physical methods for its modeling. Therefore, our approach is suitable for any two-layer networks that are represented by (1). Additionally, we use different way to extend an external cyclic structure to a complete one. We highlight a hypothesis on how the way of extension of a circle to a complete graph influences consensus time and winning rate based on simulation results and use a statistical test to verify them. The main conclusions are as follows:

- cyclic external structure always increases consensus time;
- cyclic external structure has a positive impact on winning rate;
- cyclic external structure influences consensus time by an increase of d, i.e. there is a strong linear relationship between d and consensus time. The lower d is the higher is consensus time;

a



b

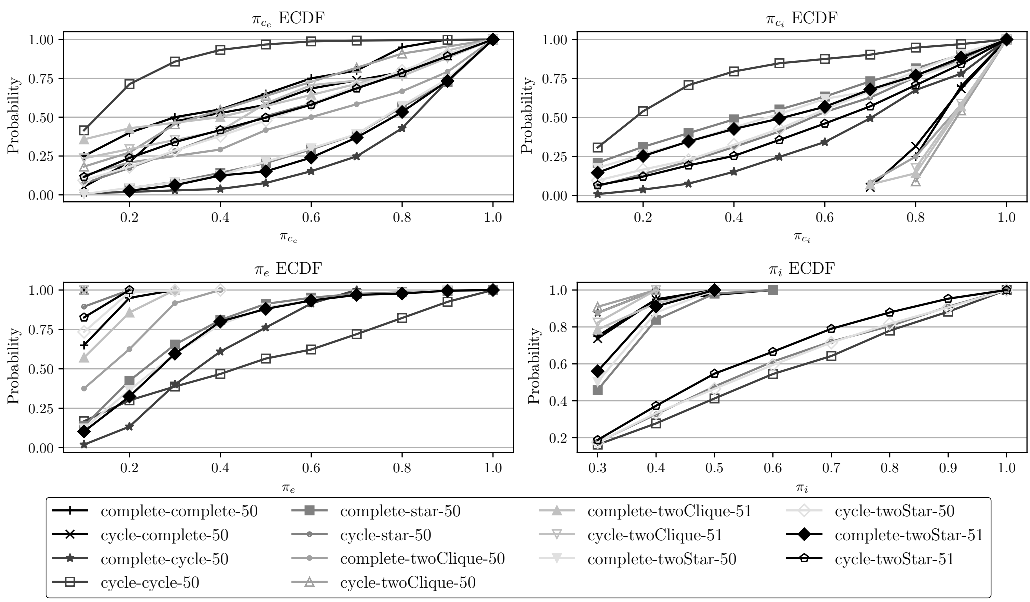


Figure 11. Empirical cumulative distribution function of consensus time with respect to parameters π_{c_e}, π_{c_i}, π_e, and π_i

a — cluster 0; *b* — cluster 1.

• the way of extension of a circle to a complete graph has a significant impact on consensus time and winning rate;

• each parameter has a different impact on consensus time, but almost has no impact on winning rate;

• the combination of parameters has a significant impact on consensus time.

We find the following developments of our work interesting: (i) analysis of other measurements (e.g., centrality) of two-layer network structures and their impact on consensus;

(ii) verification of other hypotheses based on data obtained in experiments, whether these measurements are related to consensus time and winning rate; (iii) it would also be interesting to incorporate stubbornness and redefine the condition for consensus to observe an impact of stubbornness on KPIs.

References

1. McKeehan L. A contribution to the theory of ferromagnetism. *Physical Review*, 1925, vol. 26, no. 2, pp. 274–279.

2. Holley R. A., Liggett T. M. Ergodic theorems for weakly interacting infinite systems and the voter model. *The Annals of Probability*, 1975, pp. 643–663.

3. Gastner M. T., Oborny B., Guly'as M. Consensus time in a voter model with concealed and publicly expressed opinions. *Journal of Statistical Mechanics: Theory and Experiment*, 2018, vol. 2018, no. 6, art. no. 063401.

4. Gastner M. T., Tak'acs K., Guly'as M., Szvetelszky Z., Oborny B. The impact of hypocrisy on opinion formation: A dynamic model. *PloS One*, 2019, vol. 14, no. 6, art. no. e0218729.

5. Zhao C., Parilina E. Opinion dynamics in two-layer networks with hypocrisy. *Journal of the Operations Research Society of China*, 2024, vol. 12, no. 1, pp. 109–132.

6. Noorazar H. Recent advances in opinion propagation dynamics: A 2020 survey. *The European Physical Journal Plus*, 2020, vol. 135, pp. 1–20.

7. Sznajd-Weron K., Sznajd J. Opinion evolution in closed community. *International Journal of Modern Physics C*, 2000, vol. 11, no. 06, pp. 1157–1165.

8. DeGroot M. H. Reaching a consensus. *Journal of the American Statistical Association*, 1974, vol. 69, no. 345, pp. 118–121.

9. Friedkin N. E., Johnsen E. C. Social influence and opinions. *Journal of Mathematical Sociology*, 1990, vol. 15, no. 3/4, pp. 193–206.

10. Hegselmann R., Krause U. Opinion dynamics and bounded confidence models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 2002, vol. 5, pp. 1–33.

11. Deffuant G., Neau D., Amblard F., Weisbuch G. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 2000, vol. 3, no. 01n04, pp. 87–98.

12. Parsegov S. E., Proskurnikov A. V., Tempo R., Friedkin N. E. Novel multidimensional models of opinion dynamics in social networks. *IEEE Transactions on Automatic Control*, 2016, vol. 62, no. 5, pp. 2270–2285.

13. Rogov M., Sedakov A. Coordinated influence on the opinions of social network members. *Autom Remote Control*, 2020, vol. 81, pp. 528–547.

14. Mazalov V., Parilina E. The Euler-equation approach in average-oriented opinion dynamics. *Mathematics*, 2020, vol. 8, art. no. 355.

15. Kareeva Y., Sedakov A., Zhen M. Influence in social networks with stubborn agents: From competition to bargaining. *Applied Mathematics and Computation*, 2023, vol. 444, art. no. 127790.

16. Zha Q., Kou G., Zhang H., Liang H., Chen X., Li C.-C., Dong Y. Opinion dynamics in finance and business: a literature review and research opportunities. *Financial Innovation*, 2020, vol. 6. pp. 1–22.

17. Bernardo C., Altafini C., Proskurnikov A., Vasca F. Bounded confidence opinion dynamics: A survey. *Automatica*, 2024, vol. 159, art. no. 111302.

18. Dong Y., Zhan M., Kou G., Ding Z., Liang H. A survey on the fusion process in opinion dynamics. *Information Fusion*, 2018, vol. 43, pp. 57–65.

19. Virtanen P., Gommers R., Oliphant T. E., Haberland M., Reddy T., Cournapeau D., Burovski E., Peterson P., Weckesser W., Bright J., van der Walt S. J., Brett M., Wilson J., Jarrod Millman K., Mayorov N., Nelson A. R. J., Jones E., Kern R., Larson E., Carey C. J., Polat I., Feng Y., Moore E. W., Vander Plas J., Lexalde D., Perktold J., Cimrman R., Henriksen I., Quintero E. A., Harris C. R., Archibald A. M., Ribeiro A. H., Pedregosa F., van Mulbregt P., SciPy 1.0: Fundamental algorithms for Scientific Computing in Python. *Nature Methods*, 2020, vol. 17, pp. 261–272.

20. scipy.stats.pearsonr SciPy v1.12.0 Manual, 2024. Available at: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.pearsonr.html (accessed: November 2, 2024).

21. Student. Probable error of a correlation coefficient. *Biometrika*, 1908, vol. 6, no. 2/3, pp. 302–310.

22. Simard R., L'Ecuyer P. Computing the two-sided Kolmogorov — Smirnov distribution. *Journal of Statistical Software*, 2011, vol. 39, pp. 1–18.

23. scipy.stats.kstest SciPy v1.12.0 Manual, 2024. Available at: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kstest.html (accessed: November 2, 2024).

24. Shapiro S. S., Wilk M. B. An analysis of variance test for normality (complete samples). *Biometrika*, 1965, vol. 52, no. 3/4, pp. 591–611.

25. scipy.stats.normaltest SciPy v1.12.0 Manual, 2024. Available at: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.normaltest.html (accessed: November 2, 2024).

26. Levene H. Robust tests for equality of variances. *Contributions to probability and statistics: essays in honor of Harold Hotelling*. Stanford, Stanford University Press, 1960, pp. 278–292.

27. Bartlett M. S. Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 1937, vol. 160, no. 901, pp. 268–282.

28. Fligner M. A., Killeen T. J. Distribution-free two-sample tests for scale. *Journal of the American Statistical Association*, 1976, vol. 71, no. 353, pp. 210–213.

29. scipy.stats.levene SciPy v1.12.0 Manual, 2024. Available at: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.levene.html (accessed: November 2, 2024).

30. Kruskal W. H., Wallis W. A. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 1952, vol. 47, no. 260, pp. 583–621.

31. Heiman G. W. *Understanding research methods and statistics: An integrated introduction for psychology*. Boston, Houghton Mifflin Company, 2001, 779 p.

32. Alexander R. A., Govern D. M. A new and simpler approximation for ANOVA under variance heterogeneity. *Journal of Educational Statistics*, 1994, vol. 19, no. 2, pp. 91–101.

33. scipy.stats.kruskal SciPy v1.12.0 Manual, 2024. Available at: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kruskal.html (accessed: November 2, 2024).

34. Gronlund A., Larsen K. G., Mathiasen A., Nielsen J. S., Schneider S., Song M. *Fast exact k-means, k-medians and Bregman divergence clustering in 1D*. arXiv: 1701.07204 [cs], 2018.

35. Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau A., Brucher M., Perrot M., Duchesnay E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011, vol. 12, pp. 2825–2830.

36. Rousseeuw P. J. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 1987, vol. 20, pp. 53–65.

Received: February 11, 2024.
Accepted: March 12, 2024.

Authors' information:

Chi Zhao — Postgraduate Student; st081292@student.spbu.ru

Elena M. Parilina — Dr. Sci. in Physics and Mathematics, Professor; e.parilina@spbu.ru

Анализ времени достижения консенсуса и коэффициента выигрыша в двухслойных сетях различных конфигураций в присутствии лицемерия*

Ч. Чжао, Е. М. Парилина

Санкт-Петербургский государственный университет,
Российская Федерация, 199034, Санкт-Петербург, Университетская наб., 7–9

Для цитирования: *Zhao Ch., Parilina E. M.* Analysis of consensus time and winning rate in two-layer networks with hypocrisy of different structures // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2024. Т. 20. Вып. 2. С. 170–192. https://doi.org/10.21638/spbu10.2024.204

Предложена модель общего скрытого избирателя (GCVM) микроуровня, которая сформулирована для произвольной двухслойной сети (с внутренним и внешним слоями).

* Исследование выполнено за счет гранта Российского научного фонда № 22-11-00051, https://rscf.ru/en/project/22-11-00051/

Проведена серия численных экспериментов с различными сетевыми структурами и обнаружено, что циклическая внешняя структура удлиняет время достижения консенсуса по сравнению с полной внешней структурой. Циклическая внешняя структура положительно влияет на процент выигрышей, и этот результат отличается от результата для макроскопической версии GCVM. Обсуждаются возможные причины этого различия. Кроме того, в статье подтверждена гипотеза о том, что существует сильная линейная зависимость между временем консенсуса и определенной мерой центральности в сетевой структуре. Проверено влияние каждого отдельного параметра на ключевые показатели эффективности, включая время достижения консенсуса и процент выигрышей. Оценивается влияние комбинаций параметров на ключевые показатели эффективности с использованием алгоритма K-средних. Сделан вывод, что определенные комбинации параметров могут оказать существенное влияние на время достижения консенсуса.

Ключевые слова: динамика мнений, модель избирателя, модель скрытого избирателя, обобщенная модель скрытого избирателя, частота выигрыша.

Контактная информация:

Чжао Чи — аспирант; st081292@student.spbu.ru

Парилина Елена Михайловна — д-р физ.-мат. наук, проф.; e.parilina@spbu.ru