Dynamical networks for modeling opinion dynamics in social networks

Internship report

Raphaël Truffet[1]

supervised by
Eckehard Olbrich[2] and Sven Banisch[2]

[1] École Normale Supérieure de Rennes,
Campus de Ker lann, Avenue Robert Schuman, 35170 Bruz, France

[2] Max Planck Institute for Mathematics in the Sciences,
Inselstraße 22, 04103 Leipzig, Germany

Abstract. Opinion dynamics in social networks can be modeled in different ways. During my internship, I focused on one existing model in which agents play a coordination game on a network, and they learn about opinions from social feedback, using reinforcement learning techniques. In this work, we extend this model to more than two possible opinions, and we can still observe, on geometric random graphs, the coexistence of multiple opinions. On the contrary, a consensus is reached on Erdös-Rényi models. We also add the possibility for agents to rewire in order to adapt their neighborhood instead of changing opinion. In this paper, we explore the parameter space of the model. This exploration points out different regimes in the behavior of the model which are characterized by observables such as converging time, cluster coefficient, and number of surviving opinions.

Keywords: Opinion dynamics, Games on networks, Reinforcement learning

May 15 – July 21, 2017

Table of Contents

1 Introduction

Everyone has opinions about topics such as politics, religion or even consumer products. Those opinions do not remain the same in one's life, as one is subject to several sources of influence such as friends or media. Other psychological mechanisms can also explain such opinion changes. Opinion formation and opinion dynamics have been widely studied. One of the objectives of opinion dynamics is to understand how local phenomena such as social influence or homophily can lead to some macroscopic phenomena such as consensus or polarization. Opinion dynamics only study the interactions between agents to explain the emergence of opinions. External factors, such as socio-demographic factors are not taken into account but may also be essential.

One of the first models of opinion dynamics was proposed by J.R.P French Jr [13]. It has been designed to understand complex phenomena about groups from a few simple postulates about interpersonal relations. Later work [9] described a more formal model to explain the emergence of a consensus. Most of the models are based on social influence, and lead to consensus, until new mechanisms such as bounded confidence are introduced to explain the coexistence of several opinions [14, 8]. Then, many studies [1, 19, 20, 18, 7] tried to model an opinion polarization, like S. Banisch and E. Olbrich [2] who proposed a model in which agents learn from social feedback. This model is the basis of this internship. Another important phenomenon is homophily, which is the tendency to interact with people that are similar. A simple model for homophily has been proposed by P. Holme and M.E.J Newman [15] in which agents can modify their neighborhood.

In this intership, we provide a model that combines characteristics from [2] and [15]. So, in section 2, I will first explain the original model [2], and highlight the fact that it leads to polarization. In section 3, we provide a first extension, in which there are more than two opinions. In this model, we observe that the final state depends on the initial network. Finally, in section 4, I will introduce to this model the rewiring mechanism from [15]. This adding creates a new model with several parameters, and different combinations of the parameters can lead to different behaviors of the system.

2 A game on a network

The model described in this section has been developed by S. Banisch and E. Olbrich [2]. This model can be seen as a game on network [16]. This game is a coordination game in which the objective of every agent in the network is to adapt her opinion in order to get positive reactions.

2.1 The rules of the game

The game is played on a network with N agents. They are represented by nodes of the network, and they are linked by edges. At the beginning, every agent has an opinion in the set of opinions $\mathcal{O} = \{o_{+1}, o_{-1}\}$. Fig. 1 shows agent 0's turn. In agent i's turn, the player expresses her opinion, and she receives a reaction from one of her neighbors, selected randomly (agent 1 on Fig. 1). This feedback is the only way for the agent to get information about her neighborhood, and more generally about the network. If the opinion of the neighbor is the same as the opinion expressed, then the reaction is positive (agreement) and the player gets a reward $R = +1$. Otherwise, the opinion of the neighbor is different from the opinion expressed, the reaction is negative (disagreement) and the player gets a reward $R = -1$. Then, the player perform an action: she decides to change her opinion or not. This opinion will remain until her next turn and this opinion will be expressed in the player's next turn. Note that the reward is the payoff for the move made at the end of the previous turn of the agent. The objective of every agent is to maximize the expected reward.

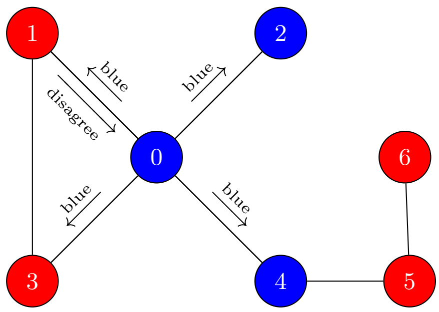


Fig. 1: One step of the game.

This game provides a partial modeling of social influence. Indeed, as the objective of the agents is to maximize the positive reactions, they might tend to have similar opinions to their neighbors. The concrete changing opinion mechanisms will come with strategies on the game.

We can describe the Nash equilibrium for this game. This equilibrium happens when every agent plays best response — no one could play better, fixing the moves of the other players. In this game, an agent plays best response when she chooses the same opinion as the majority of her neighbors. In fact, if she chooses the other opinion, she would get lower probability of getting a positive reward in her next turn. Note that an isolated agent (without any neighbor) is also considered as playing best response. In the following, we will call an agent "satisfied" if she plays best response.

2.2 A reinforcement learning approach

Reinforcement learning [21] provides solution strategies for the players. In fact, the point of reinforcement learning is also to maximize a reward. As we can see in the traditional schedule of reinforcement learning on Fig. 2, an agent receives a reward after performing an action, as well as in our game. The state of the environment can be described by the opinions of all the agents in the network. Yet, there are some differences and we have to make some approximations to use reinforcement learning techniques for our model. First, we are working with several learning agents, while there is only one in the traditional schedule of reinforcement learning. Second, every agent receives only incomplete information, only one feedback each turn. Thus, we will suppose that every agent acts as well as if she is the only one interacting with the environment, and learning from the environment. As a consequence, we suppose that every agent considers the environment constant.

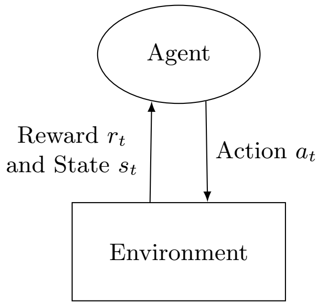


Fig. 2: Reinforcement learning schedule.

For the model, we will use a method of reinforcement learning called Q-learning. This technique consists in computing an estimate $Q(a, s)$ of the expected reward after performing action a in state s. According to our previous approximations, we will compute $Q_i(o)$, the expected reward for agent i to express opinion o, instead of $Q(a, s)$. Once Q-values are initialized, the opinion of an agent is determined by the highest Q-value: If $Q_i(o_{+1}) \geq Q_i(o_{-1})$, then agent i's opinion is o_{+1}, otherwise it is o_{-1}. Then, every time it's agent i's turn, agent i expresses this opinion, gets the corresponding reward R, and the agent finally updates her Q-value for opinion o as follow :

$$Q_i(o) \leftarrow Q_i(o) + \alpha(R - Q_i(o)) \tag{1}$$

where $0 \leq \alpha \leq 1$ is the learning rate.

This approach has two main advantages. First, this value converges to the expected reward. In fact, if R_t is the random variable of the reward received at the update number t, then the expected value of $Q_i(o)$ after update number T is:

$$\mathbb{E}(Q_i(o)) = (1-\alpha)^T Q_i(o)_{\text{initial}} + \sum_{k=0}^{T} \alpha(1-\alpha)^{T-t}\mathbb{E}(R_t) \tag{2}$$

$$= (1-\alpha)^T Q_i(o)_{\text{initial}} + \mathbb{E}(R)\left(1-(1-\alpha)^{T+1}\right) \xrightarrow[T\to\infty]{} \mathbb{E}(R) \tag{3}$$

if we consider that the R_t are independent and identically distributed random variables, which is the case with the approximations made before. The second advantage of the Q-learning update formula, is that the formula gives more weight to the latest received rewards. This is interesting because, as the constant environment is only an approximation, the Q-value contains more information about the current network.

The Q-value has also an interesting interpretation in the context of opinion dynamics. It can be seen as an evaluation of an opinion, and the update models the psychological impact of expressing this opinion. And the opinion of an agent is the opinion in which she has more confidence. While the game explains why one should change her opinion, the Q-learning gives a possible answer to the question: When should one change her opinion?

2.3 Simulation

The complete model First, in all the following experiments we initialize the Q-values at random between -1 and $+1$. In the model, the agents play in turn, and the player i is selected randomly. The player will express her opinion o, which is the opinion with the highest Q-value. Note that only the Q-value of the expressed opinion is updated. So, in order to avoid never updating some Q-values, we have to allow the agent to express the opposed opinion. This is why we introduce the exploration rate ϵ, which is the probability of expressing the opinion with the lower Q-value. Then, we select randomly the agent j that will react among the neighbors of player i. If player i has no neighbor, then the turn ends now. Otherwise, we check if the agent j's opinion is the expressed opinion o, and we compute the reaction and the corresponding reward. Then we update $Q_i(o)$ with formula 1. The turn is finished and we can select another player for the next turn. A complete step of the model is summarized in Fig. 3.

The model has been initially implemented in Matlab by S. Banisch [2], and the results are reproduced here in python. The parameters of the experiment are $N = 100$, $\alpha = 0.05$, $\epsilon = 0.1$. The initial graph is a spatial random graph [6] with parameter $r = 0.175$, which is a graph where nodes are positioned randomly in a unit square, and two nodes are connected if the distance between them is lower than r.

Results The results are presented in Fig. 4. Dispersion [10] is the variance over the distribution of convictions $\Delta Q_i = Q_i(o_{+1}) - Q_i(o_{-1})$. The support strength is the average $|\Delta Q_i|$ over the respective sets of supporters. The average opinion

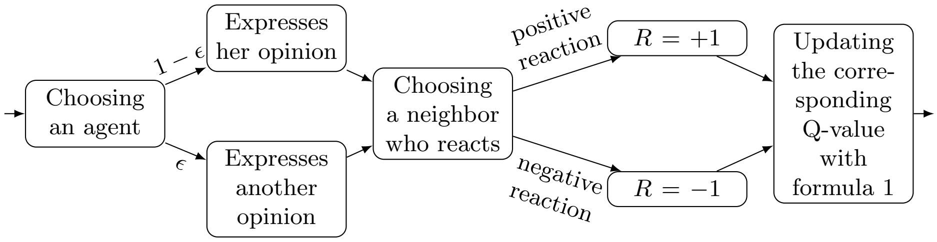


Fig. 3: One step of the model.

is the proportion of agents that express opinion o_{+1}. The dissimilarity is the polarization measure introduced in [12] whose adaptation to this model is:

$$\text{dissimiliarity} = \frac{1}{N(N-1)} \sum_{i \neq j}^{i=N, j=N} (d_{i,j} - \bar{d})^2 \tag{4}$$

where $\bar{d}$ is the average of the all the $d_{i,j}$ and

$$d_{i,j} = \frac{1}{2} \left(|Q_i(o_{+1}) - Q_j(o_{+1})| + |Q_i(o_{-1}) - Q_j(o_{-1})| \right) \tag{5}$$

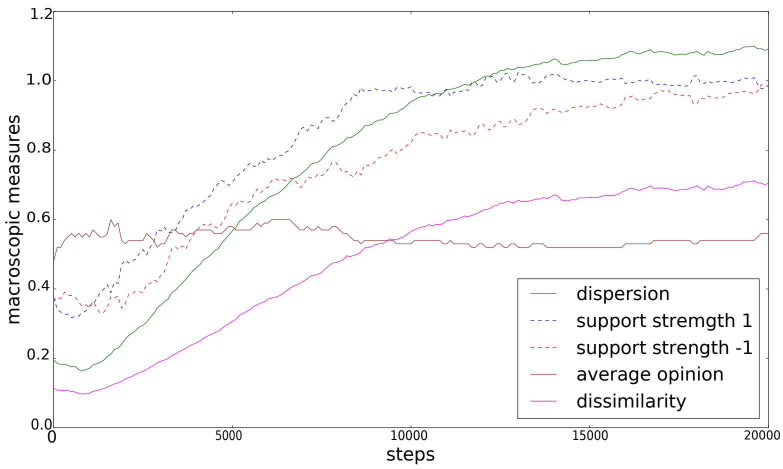


Fig. 4: Time evolution of the model.

The measures presented here reveal an opinion polarization, which is a process by which strong divergences of political opinions come about. An example of political polarization is the fundamental divide between Democrats and Republicans in the United States.

Indeed, in Fig. 4, the increase of the support strengths shows that people's beliefs in their opinion become stronger, more extreme. A high dissimilarity points out the coexistence of, on the one hand, a great number of pairs of agents with similar opinions, and on the other hand, a great number of pairs of agents with radically different opinions. Thus, dissimilarity, as well as dispersion, indicates the polarizing effect of the process. The average opinion value controls the coexistence of the two opinions in similar proportions.

3 Extension of the model to more than two opinions

The model described before was made for two opinions o_{+1} and o_{-1} only. In this section, we adapt the model to more opinions and we study its behavior on different kind of networks.

3.1 Modification of the model

For this adaptation, the schedule of Fig. 3 remains valid. There is one simple modification in the model. Instead of having 2 possible opinions, there are C possible opinions and the set of opinions becomes $\mathcal{O} = \{o_1, o_2, ..., o_C\}$. Then, every agent has C Q-values, and her opinion is determined by the highest Q-value. We also have to change the exploration mechanism as far as "the opposite opinion" does not have a sense anymore. Then, exploring now means expressing a random opinion different from the agent's opinion.

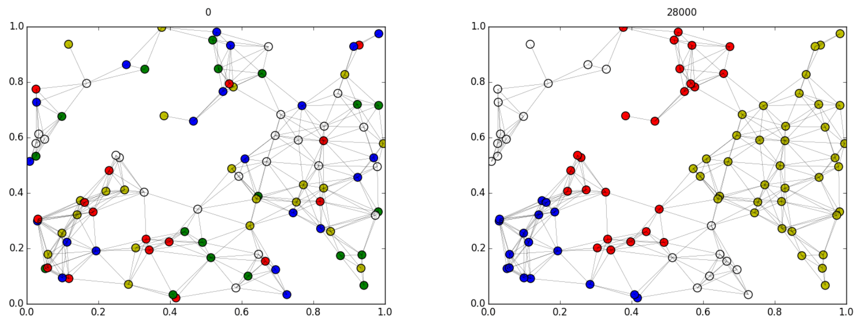
A new problem that we have to deal with, is the number of steps of an experiment. In fact, the more opinions there are in the model, the more Q-values there are, and the more steps we need to update all of them. We can use the Nash equilibrium as a stop condition.

3.2 Results over several kinds of random graphs

Spatial graph On random spatial graph, we observe the formation of communities, as we can see in Fig. 5. This is not surprising because the spatial random graphs structure tend to form geographic communities. The clustering was already observed in the two opinions model, but using more opinions allows to distinguish better the communities.

A good way to quantify the clustering is to compute the modularity. The modularity is a measure which is defined for a graph and a partition of its nodes as the difference of the proportion of edges between vertices of the same group and the same proportion if the edges were distributed at random. The modularity is given by:

$$mod(\mathcal{G}, \mathcal{P}) = \frac{1}{2m} \sum_{i,j \text{ nodes}} \left[A_{i,j} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j) \tag{6}$$



(a) Initial state (b) Final state, after 28,000 steps

Fig. 5: Model on a spatial random graph.

where m is the number of edges, k_i is the degree of node i, $A_{i,j}$ is 1 if nodes i and j are connected in the graph $\mathcal{G}$, 0 otherwise, and $\delta(c_i, c_j)$ is 1 if nodes i and j are in the same group of the partition $\mathcal{P}$, 0 otherwise. The modularity measures how good a partition of a network is. A random partition gives a modularity close to zero (it can even be negative).

In our model, the partition is given by the different opinions. We can compare the modularity value given by the model, and the one obtained with a community detection algorithm, the Louvain method [4]. On 100 random spatial graphs of $N = 100$ agents and a parameter $r = 0.175$, with $C = 100$ initial opinions, $\alpha = 0.05$ and $\epsilon = 0.1$, the average modularity is 0.58 with on average 6.39 clusters. The Louvain method on the same graphs gives partitions of on average 6.88 clusters, giving an average modularity 0.68. Even if the Louvain method gives partitions of higher modularity, community detection seems to be an interesting side effect of the model. There is even one of the 100 graphs on which the opinion dynamics model gave a partition with a higher modularity than the Louvain method.

Erdös-Rényi graph The Erdös-Rényi model [11] is one of the most common model of random graphs. It has two parameters, the number of nodes N and the probability p for two nodes to be connected.

We run the model on Erdös-Rényi graphs of $N = 100$ nodes with different values of p. Note that the parameter p could be replaced by the average degree $d = (N - 1)p$ that we can easily approximate by Np. The number of opinions is $C = 10$, the learning rate is $\alpha = 0.05$ and the exploration rate is $\epsilon = 0.1$. In fact, by running the same experiment on larger graphs, we can observe that the kind of behavior depends more on the average degree d, than the parameter p. We can observe 3 main kinds of behavior:

– For $d \leq 2$, The stop condition is quickly reached (around 50,000 steps). In the end, every opinion remains. Note that with small values of d, the graph is not connected, but even on connected components, several opinions survive. The final state can be seen on Fig. 6 for $d = 2$.
– For $2 < d < 5$, It is very long to reach the stop condition (more than 20 million steps, generally, there is only a few agents that become satisfied slowly). In the end, we can observe an opinion clustering, but only a few opinions survive.
– For $d \geq 5$, The model does not stabilize until it reaches a consensus. This consensus is reached quite fast (around 100,000 steps).

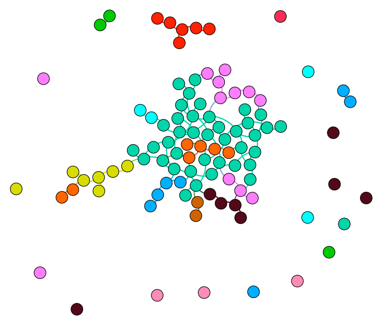


Fig. 6: Final state with average degree $d = 2$. Visualized with the software Gephi [3] using the Force Atlas 2 algorithm [17].

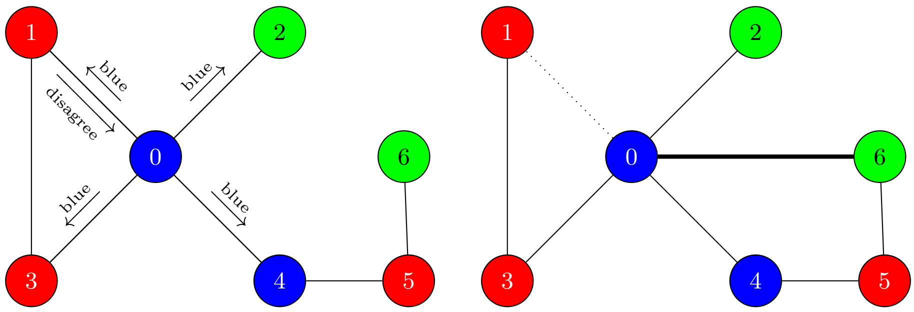
Reaching a consensus on Erdös-Rényi graphs is not surprising because, as the edges are distributed at random, the neighborhood of an agent is a sample of the whole population, so the agent will adapt her opinion to the opinion of the majority. With a low average degree, the sample may be not representative of the population, and moreover, the graph may be not connected. The intermediate value of average degree might make the equilibrium hard to reach because of a biased sample of the population. Indeed, the neighborhood of the agent is a lot influenced by the agent herself, in a way, the agent also learns from her own opinion.

4 Adding a rewiring mechanism

Until now, agents only have the possibility to adapt their opinion. Actually, an other possible approach on a network, is to let agents adapt their neighborhood to their opinion. This new mechanism represents the possibility to delete and add friends on an online social network for example. This mechanism can be modeled with the help of dynamical networks [5].

4.1 The rewiring mechanism

This mechanism is inpired by P. Holme and M.E.J Newman [15]. We introduce to our model a rewiring rate φ. The possibility to rewire appears when an agent i receives a negative reward from an agent j. This agent will, with probability $1 - \varphi$, update her Q-value as it was done until now, and with probability φ, delete the edge between i and j, and create a new edge between i and a random new neighbor (see Fig. 7).



(a) Receiving a negative reward (b) Rewiring

Fig. 7: Rewiring mechanism.

The schedule in Fig. 3 will be augmented with Fig. 8. The stop condition is no more a Nash equilibrium, a "satisfied" agent could still change some neighbors to have an even better neighborhood. With the rewiring mechanism, the Nash equilibrium is more complex to identify, because there are several reasons to explain that an agent is playing best response:

- She has the same opinion as all her neighbors.
- She is connected to every agent that has the same opinion as her, and this opinion is the opinion of the majority in her neighborhood. Note that if there is in the network an opinion that affects more agents, the agent could still change her opinion and rewire to connect to a maximum of agents with this opinion to improve her expected reward, but this would take several steps.

So, we still consider as satisfied an agent with the same opinion as the majority of her neighbors, and stop the simulations when all agents are satisfied. This stop condition is more convenient, because, it is simpler and, in the following the rewiring rate may be 0, and the final state is still close to a Nash equilibrium.

In this model, as the total number of edges remains constant, then the average degree and the density does not evolve neither. In order to observe the evolution of the connections in the graph, we can compute the average local clustering coefficient, which express the tendency of the graph to form triangles.

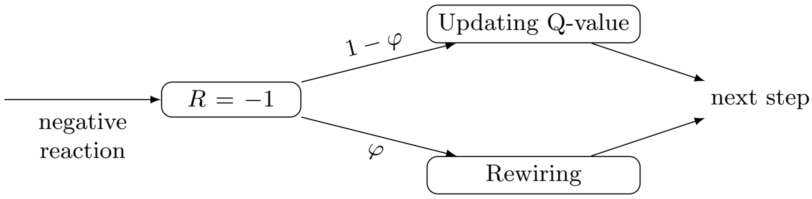


Fig. 8: Rewiring schedule.

This average clustering coefficient is defined in [22] as follow :

$$acc = \frac{1}{N} \sum_{i \text{ node}} C_i \tag{7}$$

where the clustering coefficient of a node is :

$$C_i = \frac{\text{number of pairs of neighbors connected}}{\text{degree}_i \times (\text{degree}_i - 1)} \tag{8}$$

4.2 Evolution in time

The first approach to study the new model is to observe the evolution of several measures in time. On Fig. 9, we can follow the evolution of the modularity, the average clustering coefficient, the proportion of agent satisfied as defined in section 2.1, and the proportion of opinions remaining compared to the total number of possible opinions. The experiment has been run for an Erdös-Rényi graph of $N = 1000$ agents and a density $p = 0.1$, $C = 100$ possible opinions, a learning rate $\alpha = 0.05$, an exploration rate $\epsilon = 0.1$ and a rewiring rate $\varphi = 0.5$. We can observe that the measures do not evolve a lot until a sudden transition. This transition is quick compared to the time elapsed before. After that, all the measures stabilize to a final value. Note that we did not stop the simulation when all agents are satisfied in this experiment.

The final state can be observed in Fig 10. The graph presented has been visualized with the software Gephi using the Force Atlas 2 algorithm. In the end, the network contains several connected components of several sizes. On every component, a consensus is reached.

As the system stabilized after a transition, in what follows, we only get interested in the final values of the measures.

4.3 Exploration of the parameters space

With $N = 100$, $C = 10$, $\epsilon = 0.1$ on an Erdös-Rényi graph with parameter $p = 0.1$, we ran 100 experiments over 31×31 pairs of values of (α, φ). As the the result of an experiment is not deterministic, averaging over these 100 experiments gives reliable estimates. The mean values of several measures are presented on Fig. 11.

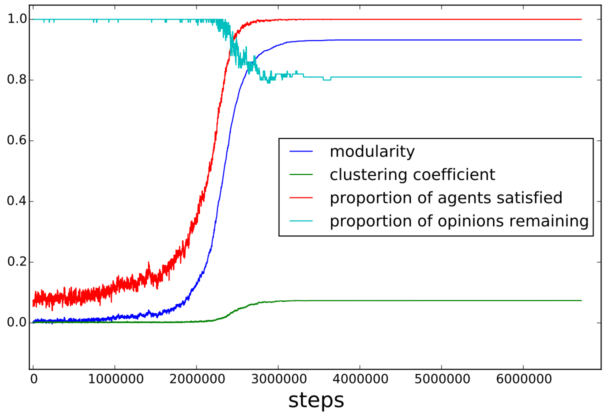


Fig. 9: Time evolution of the model.

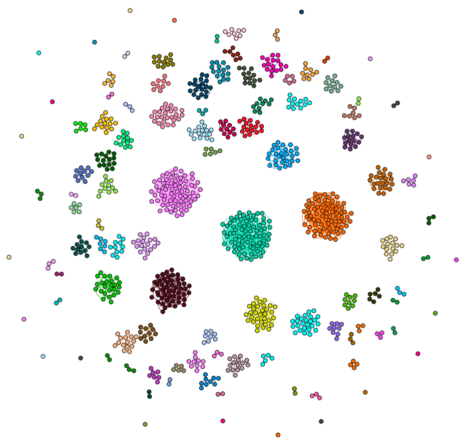
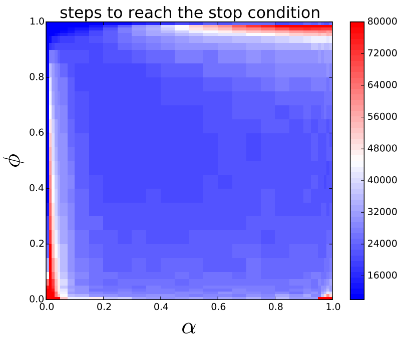


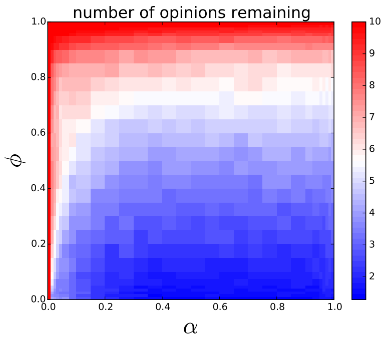
Fig. 10: Final state.

Fig. 11a shows the time to reach the stop condition. It can be interpreted as how efficient the parameters are to satisfy the agents — in the sense that how fast the stop condition is reached. For example, it does not seem efficient to set both α and φ at a low value, or both at a high value. The most efficient regime seems to be a high φ with a low α. The line $\varphi = 1$ also seems interesting. Finally, the fact that the line $\varphi = 1 - \alpha$ looks efficient is intuitive: The more you can change your opinion, the less you need to change your friends.

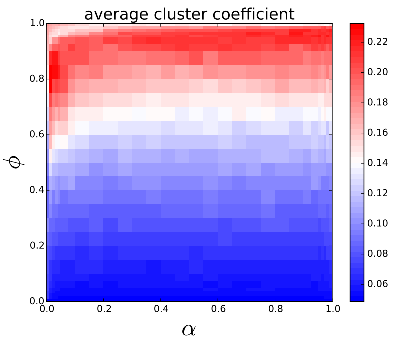
But even if Fig. 11a shows several possible regimes, Fig. 11b and 11d show that those regimes lead to different final states. For example, the most efficient regime described earlier and the line $\varphi = 1$, lead to a final state in which many



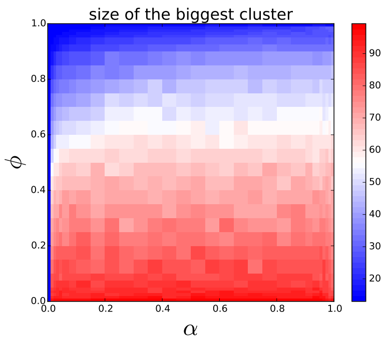
(a) Duration of the experiment.



(b) Survival of opinions.



(c) Tendancy to form triangles.



(d) Size of the biggest cluster.

Fig. 11: Observables over the parameter space.

opinions, forming several clusters of a similar size. On the line $\varphi = 1 - \alpha$, the higher α is, the more one opinion is supported, implying the predominance of one cluster. The final clustering coefficients presented in Fig. 11c also reveal different ending structures of network depending on the regime. The modularity is not presented here because, as the number of cluster varies a lot depending on the parameters, the modularity measure may not be relevant.

In Fig. 12, we show the number of clusters of different sizes in the final state of the 100 experiments for 3 values of φ and $\alpha = 0.05$. This result is very similar to what is observed by P. Holme and M.E.J Newman [15] in their model. We can see a qualitative change from a regime with no giant community to one with a giant community. At an intermediate value of φ, we find a distribution of community sizes that appears to follow a power law. P. Holme and M.E.J Newman show that this intermediate value is a critical value that separates two regimes in which a giant component appears or not. As we can guess from Fig. 11b and 12, α has an influence on this critical φ-value for $\alpha \leq 0.2$.

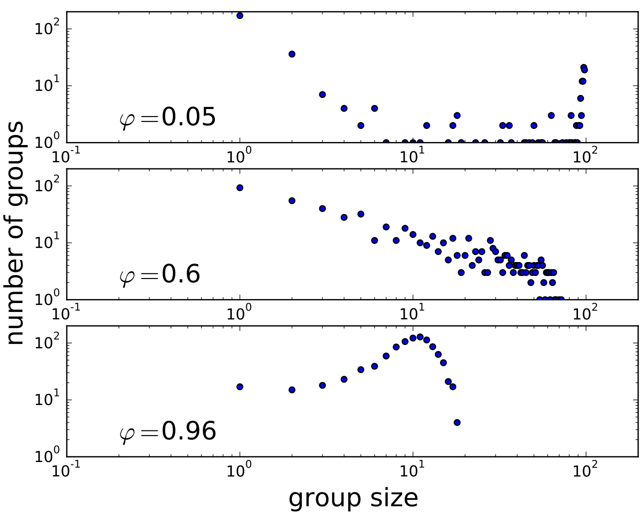


Fig. 12: Community sizes for 100 experiments with $\alpha = 0.05$.

5 Conclusion

We provided, during this internship, two extensions to an existing model of opinion dynamics [2]. The first extension provides a model for several opinions instead of two. The increase of the number of opinions does not impede to reach an equilibrium. The equilibrium depends on the initial network. On random spatial graphs, the equilibrium is characterized by an opinion clustering of the network. On Erdös-Rényi graphs, if the average degree is big enough, we can observe the emergence of a consensus.

The second extension allows agents to rewire. This possibility leads to the formation of disconnected clusters. Depending on the parameters of the model, we can observe several kinds of regimes, in which a giant component appears or not. In fact, the lower the rewiring rate is, the more one opinion dominates. With a high rewiring rate, the final clusters have similar sizes. Between these two regimes, there is a critical value of the rewiring rate for which the distribution of the sizes of the clusters follows a power law. This critical value depends on the learning rate.

Yet, these results are obtained with many parameters of the model fixed. Even if we ran some experiments with different values for those parameters – exploration rate, number of possible opinions, number of agents, graph density – the exploration is not exhaustive. At least, the few experiments that have been run show the same kind of behavior, but with different critical values. A further study would be necessary to highlight a correlation between those parameters and the observations.

The model provided here and tested on simple random graphs, could also give interesting results over more complex structures. For example, one can wonder what the behavior of the system is on a configuration model, or on real data networks. The model without rewiring might be more sensitive to the initial network configuration because the rewiring mechanism leads to a reconfiguration of the network, so the initial configuration does not persist for a long time.

It would also be interesting to use the Q-learning approach not only for choosing an opinion, but also to give an evaluation of the rewiring mechanism as an action performed by the agent.

References

[1] Delia Baldassarri and Peter Bearman. "Dynamics of political polarization". In: *American sociological review* 72.5 (2007), pp. 784–811.

[2] Sven Banisch and Eckehard Olbrich. "Opinion Polarization by Learning from Social Feedback". In: *CoRR* abs/1704.02890 (2017).

[3] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. *Gephi: An Open Source Software for Exploring and Manipulating Networks.* 2009.

[4] Vincent D Blondel et al. "Fast unfolding of communities in large networks". In: *Journal of Statistical Mechanics: Theory and Experiment* 2008.10 (2008), P10008.

[5] K.M. Carley. "Dynamic network analysis". In: *Dynamic social network modeling and analysis: Workshop summary and papers.* 2003, pp. 133–145.

[6] Jesper Dall and Michael Christensen. "Random geometric graphs". In: *Physical Review E* 66.1 (2002), p. 016121.

[7] P. Dandekar, A. Goel, and D. T. Lee. "Biased assimilation, homophily, and the dynamics of polarization". In: *Proc. Natl. Acad. Sci. U.S.A.* 110.15 (Apr. 2013), pp. 5791–5796.

[8] Guillaume Deffuant et al. "Mixing beliefs among interacting agents". In: *Advances in Complex Systems* 03.01n04 (2000), pp. 87–98.

[9] Morris H DeGroot. "Reaching a consensus". In: *Journal of the American Statistical Association* 69.345 (1974), pp. 118–121.

[10] Paul DiMaggio, John Evans, and Bethany Bryson. "Have American's Social Attitudes Become More Polarized?" In: *American Journal of Sociology* 102.3 (1996), pp. 690–755.

[11] Paul Erdős and Alfréd Rényi. "On the Evolution of Random Graphs". In: *PUBLICATION OF THE MATHEMATICAL INSTITUTE OF THE HUNGARIAN ACADEMY OF SCIENCES.* 1960, pp. 17–61.

[12] Andreas Flache and Michael W. Macy. "Small Worlds and Cultural Polarization". In: *The Journal of Mathematical Sociology* 35.1-3 (2011), pp. 146–176.

[13] John RP French Jr. "A formal theory of social power." In: *Psychological review* 63.3 (1956), p. 181.

[14] Rainer Hegselmann and Ulrich Krause. "Opinion Dynamics and Bounded Confidence, Models, Analysis and Simulation". In: *Journal of Artificial Societies and Social Simulation* 5.3 (2002), p. 2.

[15] Petter Holme and Mark EJ Newman. "Nonequilibrium phase transition in the coevolution of networks and opinions". In: *Physical Review E* 74.5 (2006), p. 056108.

[16] Matthew O. Jackson and Yves Zenou. "Games on Networks". In: vol. 4. Handbook of Game Theory with Economic Applications. Elsevier, 2015, pp. 95–163.

[17] Mathieu Jacomy et al. "ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software". In: *PLOS ONE* 9.6 (June 2014), pp. 1–12.

[18] Noah P Mark. “Culture and competition: Homophily and distancing explanations for cultural niches”. In: *American sociological review* 68.3 (2003), pp. 319–345.

[19] Michael Mäs and Lukas Bischofberger. “Will the Personalization of Online Social Networks Foster Opinion Polarization?” In: *Available at SSRN 2553436* (2015).

[20] Michael Mäs and Andreas Flache. “Differentiation without distancing. Explaining bi-polarization of opinions without negative influence”. In: *PloS one* 8.11 (2013), e74516.

[21] Richard S. Sutton and Andrew G. Barto. *Introduction to Reinforcement Learning*. 1st. Cambridge, MA, USA: MIT Press, 1998.

[22] Duncan J. Watts and Steven H. Strogatz. “Collective dynamics of /‘small-world/’ networks”. In: *Nature* 393.6684 (June 1998), pp. 440–442.