Article

Community Formation as a Byproduct of a Recommendation System: A Simulation Model for Bubble Formation in Social Media

Franco Bagnoli [1,2,*], Guido de Bonfioli Cavalcabo' [1], Banedetto Casu [1] and Andrea Guazzini [3]

[1] Department of Physics and Astronomy and CSDC, University of Florence, Via G. Sansone 1, 50019 Sesto Fiorentino, Italy; guido.debonfiolicavalcabo@stud.unifi.it (G.d.B.C.); benedetto.casu@stud.unifi.it (B.C.)

[2] INFN, sect. Florence, Via G. Sansone 1, 50019 Sesto Fiorentino, Italy

[3] Department of Education, Languages, Intercultures, Literatures and Psychology and CSDC, University of Florence, Via Laura 48, 50121 Firenze, Italy; andrea.guazzini@unifi.it

[*] Correspondence: franco.bagnoli@unifi.it

Abstract: We investigate the problem of the formation of communities of users that selectively exchange messages among them in a simulated environment. This closed community can be seen as the prototype of the bubble effect, i.e., the isolation of individuals from other communities. We develop a computational model of a society, where each individual is represented as a simple neural network (a perceptron), under the influence of a recommendation system that honestly forward messages (posts) to other individuals that in the past appreciated previous messages from the sender, i.e., that showed a certain degree of affinity. This dynamical affinity database determines the interaction network. We start from a set of individuals with random preferences (factors), so that at the beginning, there is no community structure at all. We show that the simple effect of the recommendation system is not sufficient to induce the isolation of communities, even when the database of user–user affinity is based on a small sample of initial messages, subject to small-sampling fluctuations. On the contrary, when the simulated individuals evolve their internal factors accordingly with the received messages, communities can emerge. This emergence is stronger the slower the evolution of individuals, while immediate convergence favors to the breakdown of the system in smaller communities. In any case, the final communities are strongly dependent on the sequence of messages, since one can get different final communities starting from the same initial distribution of users' factors, changing only the order of users emitting messages. In other words, the main outcome of our investigation is that the bubble formation depends on users' evolution and is strongly dependent on early interactions.

Keywords: community formation; filter bubble; echo chamber; recommendation systems; social media; agent-based simulations

1. Introduction

Starting from the widespread use of the internet, and especially with the worldwide diffusion of social media, like Facebook, concerns about the effect of recommendation systems started to rise [1,2].

Since social media contains a huge amount of information, a simple search retrieves too many results to display them all, therefore only a small amount of those are selected. Among them, there are the posts from people and the pages that the user actively selected to follow, but also those recommended by the system's algorithm.

Recommendation systems can be broadly classified into three categories accordingly to the type of algorithms used: social engineering-based, content-based, and collaborative filtering [3,4].

Social engineering methods are based on the weaknesses of our decision systems [5] that are essentially related to the dual process theory, one being an implicit (automatic), unconscious process, and the other an explicit (controlled), conscious one. The first process is based on heuristics, i.e., "rules of thumb" based on previous experiences and easy to retrieve, while the second consists in rational thinking. The first process is preferred unless there is a strong reason to check the rationality of the behavior, and is strongly promoted in situations of danger, stress, and/or bounded time limit. These situations are typically enforced by vendors and scammers. Social engineering often consists in the participation in active discussions, interviews, focus groups, etc. These two methods require a certain amount of human work.

Content-based systems use the classification of the characteristics of personal information disclosed by users.

An improvement of this latter method is given by collaborative filtering, which is based on the harnesses of information that can be gathered from users' communications (e.g., browsing history or other pieces of interaction). This method extracts knowledge from transactions (that will be addressed in the following as "posts", having in mind an exchange of messages on social media), and stores them in a database, to be confronted with the ones of other users. Using the stored information on the reading or active evaluation ("likes") of users on other users' posts, collaborative filtering can extrapolate information regarding the similarities among those users and recommend unread posts only of those which the system believes to be similar. The same techniques obviously apply to e-commerce and advertising.

One possible drawback of this kind of recommendation system is the formation of filter bubbles [6,7] or echo chambers [8,9], i.e., the formation of intellectually closed circles of users that exchange information only about selected topics or share the same opinion. It was pointed out that an increased separation may lead to the exacerbation of conflicts and radicalism [10].

Many critiques to this extreme point of view exist [11], and also various suggestions on how to break those filters and pop the bubble [12], however, a secondary effect of collaborative filtering might be the formation of "artificial" communities; groups where the "true" affinity among peoples is not reflected, and they happen to be in the same bubble just as a byproduct of the recommendation system.

There were experiments trying to induce the formation of communities and bubbles on YouTube [13] and developing metrics to characterize the "missing information" not received by a user due to the recommendation system [14].

However, it is not clear from field experiments what are the users' characteristics that favor the formation of bubbles and communities. Is it sufficient to filter out information, or does the mechanism only work when users modify their preferences? Do the final communities reflect an initial polarization of users, or are they determined by the random fluctuation of message exchanges?

These aspects can play a fundamental role in the social dynamics. For instance, if bubble formation can appear even when users do not change their minds, a "restart" of the recommendation system (or a switch to another social media) could bring to the formation of completely different communities, even with the same users. On the contrary, if the formation of communities is strongly dependent on users' evolution, the role of stubborns, i.e., of people never changing their mind, may reveal to be crucial.

Finally, if the emerged communities do not reflect any initial polarization, but are the result of random fluctuations, strong consequences on the interpretation of the social factors that promote radicalization are expected: maybe the efforts should focus on individual formation (propensity towards diversity) or on forcing modifications on social media, rather than on conveying proactive messages.

Since it is impossible to play "sliding doors" experiments on real social media, we develop a simulated environment, trying to capture the essence of user cognitive dynamics and modeling a rough recommendation system that exploits its knowledge on expressed

preferences to determine user-user affinity and modify their future interactions. The role of numerical, agent-based experiments is also emphasized in Ref. [15].

The main goal of a simulated environment is that of reproducing an effect that was observed in a real situation by reducing at minimum the "ingredients" of the model, to focus out the essential element that originates a given behavior.

Users are modeled as simple perceptrons [16], i.e., they are characterized by a certain number of internal factors, which determine the opinion over a given message by the match between them and the characteristics of the message itself [17].

A perceptron is a simple linear classifier, i.e., a classification algorithm that makes its predictions based on a linear predictor function combining a set of weights with a feature vector [18].

The perceptron modeling is coherent with the factorial analysis [19–21], which is a method used to describe variability among observed, correlated variables in terms of a potentially lower number of unobserved and independent variables called factors.

Indeed, the idea of representing people as vectors is quite old [22]. The use of perceptrons implies the use of a learning algorithm which allows us to evolve the system evolving every single users after they had an interaction.

Since a single post cannot express an opinion on all possible topics, we assume that when people emit those posts they are expressing only a subset of their inner factors. The post is then evaluated by other users who, through the perceptron key-and-lock mechanism, express their opinion on it.

By means of the set of expressed opinions, it is possible to approximate the similarity among users [17,23,24]. However, we are here interested in studying what happens when users are not exposed to all messages or to a random sample of them, but rather only to those recommended by the system based on a previous (limited) sampling.

Our simulated recommendation system works by recording the opinion of users on received messages (given by the overlap between the factors of receiving users and those contained in the message, originating from emitter's ones). Based on such an affinity database, the recommendation system selects the fraction of users to which a message is propagated. So the network between users (i.e., who receives a certain message) is dynamically determined by the affinity and has no pre-established structure.

We begin considering people immutable (they do not change their opinions), and in a second moment we let them evolve in a way that reduces their cognitive dissonance, i.e., "aligning" with the opinions expressed in the received messages.

Our model contains many elements of originality. In most cases, simulations are based on agent-based models (ABM), in which the dynamics directly affects opinions [25–28], not the factors determining them. When factors are explicitly included, as in [29], they are limited to two, and (for the limitations of the used Netlogo platform) with a very small number of users.

The recommendation system is rarely explicitly implemented, resorting to a bounded confidence approach [25–29]. There are cases in which there is recommendation system based on a database of opinions, as in [30], but in this case the interaction considers only on the dichotomy between core and peripheral interests.

Finally, we were unable to find other cases in which the replica symetry breaking ("sliding doors" experiment) was investigated in the context of bubble/community formation.

In Section 2 we illustrate mathematically our model and describe the simulation scheme. The results of simulations are presented in Section , firstly examining the case in which users do not change their minds, and then by allowing them to "align" their factors with the received messages.

We then investigate what happens when the same initial set of users emit and receive messages in a different order, i.e., simulating a "sliding doors" (replica symmetry breaking) experiment.

Conclusions are drawn in the last section.

2. The Model

Users are represented as vectors of L factors that represent the possible topics on which they can express an opinion, so that user i corresponds to $U_i = \left(u_i^{(1)}, u_i^{(2)}, \ldots, u_i^{(L)} \right)$, with $-1 \leq u_i^{(k)} \leq 1$, $1 \leq k \leq L$ and $1 \leq i \leq N$, where N is the total number of users. Similarly, since posts might contain opinion on a range of those topics, they are represented as vectors of L components as well: $P_n = \left(p_n^{(1)}, p_n^{(2)}, \ldots, p_n^{(L)} \right)$ with the same range $-1 \leq p_n^{(k)} \leq 1$ but with $1 \leq n \leq M$, where M is the total number of messages.

The opinion Ω_{in} of user i about post n is computed as

$$\Omega_{in} = \tanh\left(\frac{\beta}{L} \sum_k u_i^{(k)} p_n^{(k)} \right).$$

The factor β modulates the nonlinearities of the system. If β is large, the system is basically a Heaviside function, if β is small the system is essentially linear. In the following, we shall deal only with the linear case:

$$\Omega_{in} = \frac{1}{L} \sum_k u_i^{(k)} p_n^{(k)}.$$

Given the matrix U of users' factors, such that $U_{ik} = u_i^{(k)}$ and a set of M messages such that a single post can be written as $P_{nk} = p_n^{(k)}$, we compute the opinion matrix Ω,

$$\Omega = \frac{1}{L} U P^T.$$

2.1. Recommender Systems

The systems starts by computing the Fisher correlation matrix between user i and j by considering the expressed opinions Ω_{in} and Ω_{jn} as

$$C_{ij} = \frac{1}{M-1} \sum_{n=1}^{M} \left(\frac{\Omega_{in} - E(\Omega_i)}{\sigma(\Omega_i)} \right) \left(\frac{\Omega_{jn} - E(\Omega_j)}{\sigma(\Omega_j)} \right),$$

where M is the number of messages present in the database, $E(\Omega_i)$ is the average of the M-vector Ω_i and $\sigma(\Omega_i)$ is the corresponding standard deviation.

Once user i emits a post p_n at "time" n, the recommender system looks for the ensemble $V(i)$ of all people j, that have correlation with i greater than a threshold τ,

$$V(i) = \{ j : C_{ij} > \tau \}$$

and exposes them to the message. The corresponding opinions are recorded in the data series. All other people are assigned opinion zero (neutral).

2.2. The Procedure

We begin the process by generating N users with random factors $u_i^{(k)}$, then we initialize the database using the opinion of all users on a number M_0 of messages emitted by random users, delivered to the whole population.

After that, for a number of steps M, one user is drawn at random, and we extract a number of factors (dimensions of the message) by choosing each of them with probability f, so that in average a message covers fL topics. Users therefore emit messages that reflect their opinion for the chosen number of dimensions. Our users and their messages are honest and there is no possibility of distortion. The number of messages M is chosen so as to have an asymptotic state, i.e., nothing changes if doubling M.

We then propose that message to all people that have a correlation with the emitter greater than τ, modifying the database accordingly, as explained in the previous section.

If we let evolve people's opinions, we use a parameter ε that models memory: the receiver factors tend to align to the ones appearing in the received message for a percentage ε, in a way similar to what happens for one user in the Deffuant model [25,26].

This evolution take place updating, due to posts p whose factors (each one with probability f) are $p_k = U_{jk}$, the factor k of user i, U'_{ik},

$$\begin{cases} U_{ik} + \varepsilon(p_k - U_{ik}) & \text{if } p_k = U_{jk}; \\ U_{ik} & \text{otherwise.} \end{cases}$$

2.3. Community Detection

We used the MATLAB function `linkage` to produce dendrograms of the overlap among individuals. This function starts by assigning at first a cluster to each item. Then, it proceeds by finding the two nearest clusters and replacing them both with a cluster located at the average distance between them. The point where two clusters join is marked at a height proportional to the distance between them. This procedure is repeated until there is only one cluster left, see Figure 1a.
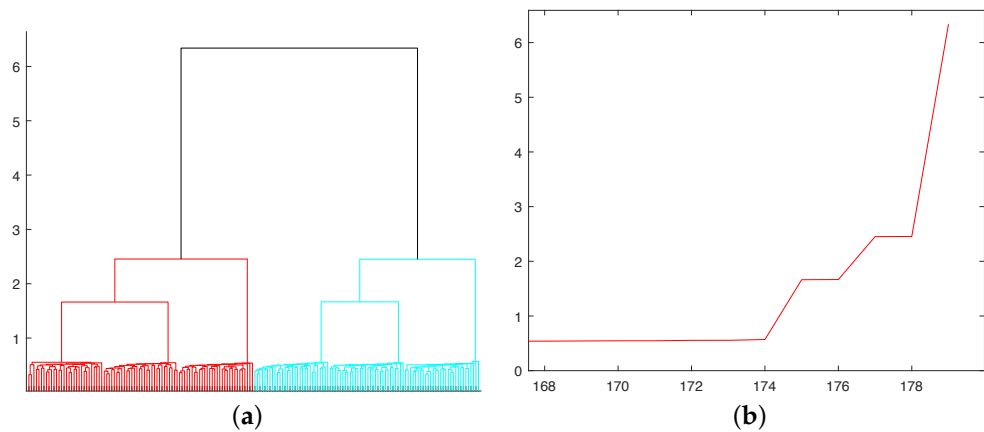

Figure 1. Community detection algorithm with a tailored example with half positive values of opinions and half with negative values. (**a**) Linkage mechanism: as we can see we have two community (red and blue) formed by a few other groups connecting users. (**b**) Distribution of jumps: in this graph the distance between opinions of users or group of users is reported, as we can see largest jump could be used as an indicator of presence of a community.

By plotting the sequence of heights of the joining point, Figure 1b, one can see "jumps" corresponding to the merging of two clusters. When there is no clear community structure, the joining occurs almost continuously and there are no large jumps, as happens in Figure 1a at the beginning of the clustering procedure. But when people are grouped into well-separated communities, the distance between two clusters is larger and therefore there is a large jump in the distance, as happens for the 6 communities in Figure 1a, corresponding to the jumps in Figure 1b. We therefore use the largest jump as an indicator of the presence of a community structure.

The maximum jump indicator is sensible to statistics: by increasing the communication threshold this indicator is decreasing, since the database is less populated, as shown inf Figure 1b. The same effect is present for different values of f and also by simply increasing the number of initial messages Q_0, and is reflected in the apparent presence of communities even in the absence of evolution.

Since users are generated at random, no community structure is present in the real user overlap O,

$$O_{ij} = \frac{1}{L} \sum_{k=1}^{L} u_i^{(k)} u_j^{(k)}.$$

3. Results

We performed simulations with various values of L (number of factors), N, (number of users), M_0 (number of initial messages), M (number of messages), f (fraction of factors in emitted messages) and τ (threshold for selecting which user is receiving a recommendation). In the following, we shall report the results of experiments concerning parameters that determine a change in the behavior of the system [2,9,31]. When not otherwise stated, the values chosen for the simulations are already in the "thermodynamic" limit, i.e., results do not change by increasing them. In particular, the influence of f, at least for $f \geq 0.2$, is negligible and it will not be reported.

In the following, the unit of time corresponds to the delivering of one message.

3.1. No Factor Evolution

The first case we analyze is the one without the evolution of users opinions ($\varepsilon = 0$).

The time evolution of the structure is extremely slow. As reported in Figure 2a, there are often early big fluctuations in the community indicator, which slowly converges to a final value. The convergence is slower for large values of τ, since, in this case, the message is sent only to a limited number of people.

In the case of a small number N of users and large number L of factors, spurious communities can appear because the small number of items will be sparse in a space of L dimensions. We noted that we obtained very similar results if the number of users is scaled with the number of factors.
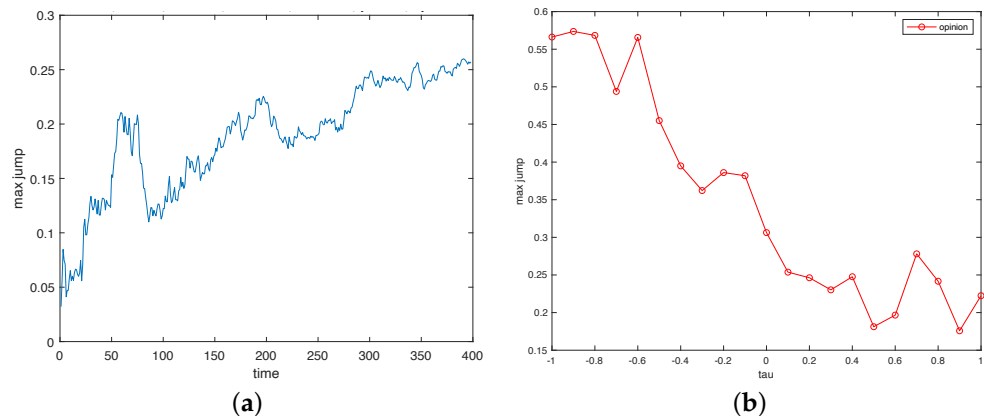

(**a**)

(**b**)

Figure 2. (**a**) Time evolution of the community indicator for a simulation with $\varepsilon = 0$ $L = 40$, $N = 100$, $M_0 = 200$, $\tau = 0.5$, $f = 0.4$, $\varepsilon = 0$, $M = 40{,}000$. The unit of time corresponds to the sending/receiving of one message, and the time scale (number of messages sent) is in unit of 10^2. (**b**) Dependence of the maximum jump indicator on the threshold τ in a system without evolution ($\varepsilon = 0$). The other parameters are $L = 20$, $N = 100$, $M_0 = 200$, $\tau = 0.5$, $f = 0.4$, $M = 10{,}000$.

On the contrary, when there is a large number of users with respect to factors, no community structures appear in the users overlap (Figures 3a and 4a).
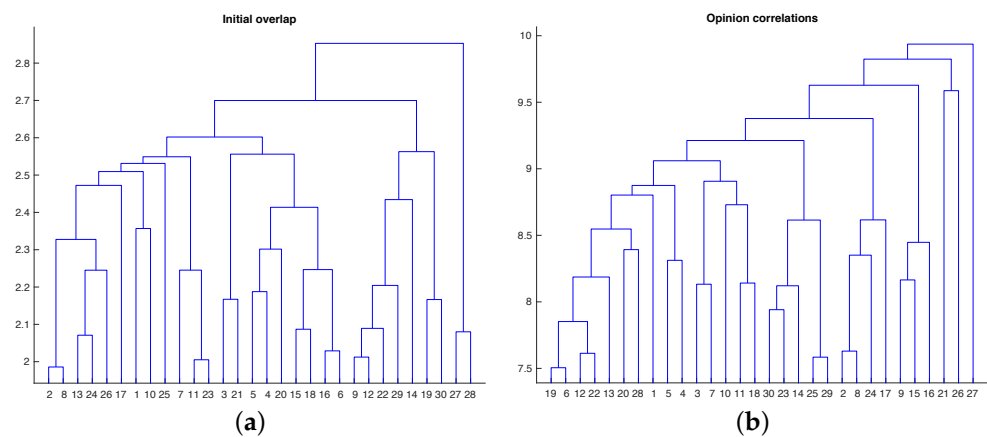
Figure 3. Dendrograms of relations between users. Distance between clusters is reported as difference between height. Value of parameters are: $L = 10$, $M_0 = 200$, $\tau = 0$, $f = 0.4$, $\varepsilon = 0$, $N = 200$, $M = 20{,}000$. (**a**) Here, we can see user overlap, or how big differences are between opinions of single users, while in (**b**) we can see opinion overlap. No evident "jump" is seen, which is reasonable since users are randomly generated and selection threshold (τ) is zero.
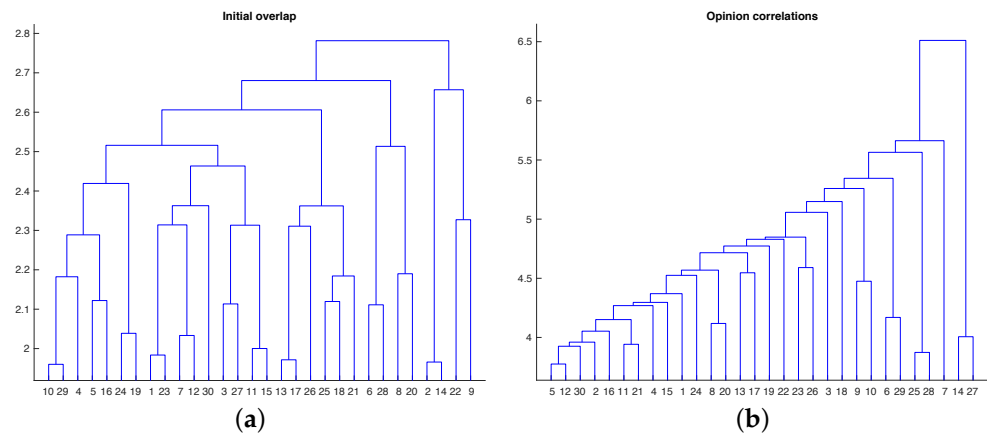

Figure 4. Community detection with $L = 10$, $M_0 = 200$, $\tau = 0.4$, $f = 0.4$, $\varepsilon = 0$, $N = 200$, $M = 20{,}000$. (**a**) user overlap, (**b**) opinion overlap. As we can see an higher threshold causes the appearance of a "jump" in coalescence of clusters, but this is an effect due to statistics, as explained in text.

By waiting a sufficiently long time, the cluster structure emerging from the opinion database is the same as the real user overlap, as shown in Figure 3. This is consistent with the fact that users overlap can be recovered by examining the opinions expressed on the messages [17].

By changing the selection factor τ for a fixed evolution time M we see a decreasing of the clustering indicator (maximum jump), as reported in Figure 2b. This is due to the fact that for large values of τ the messages are sent only to neighboring users (in the factor space). So the initial database structure, which is given by the small number of initial messages M_0 remains almost frozen, until, slowly, users express opinions on messages sent by intermediate users updating the database and making the opinion correlation slowly converge to the real overlap O.

3.2. Evolution of Factors

If we let user opinion vary with time ($\varepsilon > 0$), the "communities" in the opinion database that determine the selection of messages induce the formation of real communities in the final user overlap, as we can see in Figures 5 and 6. The faster the adherence to local conformism, the larger the number of resulting of communities, and therefore the smaller the maximum jump.
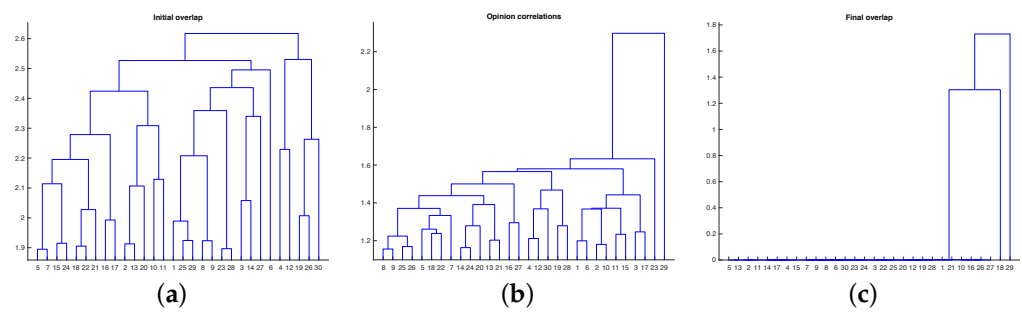
Figure 5. Community detection with $L = 10$, $M_0 = 200$, $\tau = 0.4$, $f = 0.4$, $\varepsilon = 0.3$, $N = 200$, $M = 10{,}000$. (**a**) Initial overlap between users, (**b**) overlap between opinions, (**c**) final user overlap after letting system evolve. A nonzero value of ε in an evolving system completely separates population.
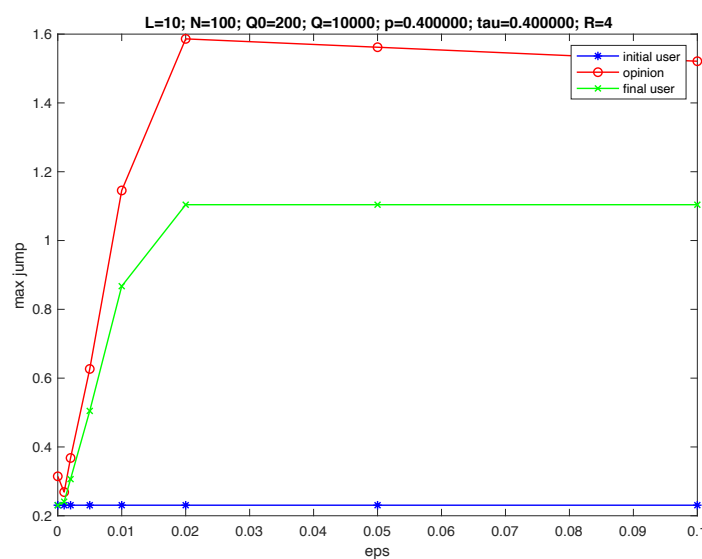

Figure 6. Simulation with: $L = 10$, $M_0 = 200$, $\tau = 0.4$, $f = 0.4$, $N = 200$, $M = 10{,}000$. Largest jump in dendrogram as a function of ε. We can note that as soon as $\varepsilon > 0$ a strong sign of community formation appear both in opinions and in final factors of users. Community structure weakens a bit by increasing ε since in this case users form more numerous but smaller communities.

3.3. Replica Breaking

The most interesting effect concerns the final structure of the users' factors obtained by changing the sequence of messages. We repeated the evolution process with the same initial factors for R replicas, but every time with different sequences of random extraction of users emitting messages. We then measured the distance D_{1r} between replica 1 and r by averaging the distance of evolved users in replica r, $U_{ik}^{(r)}$ from the same users in the first replica $U_{ik}^{(1)}$, and averaging over all replicas,

$$D = \frac{1}{NR \cdot L} \sum_{r=1}^{R} \left(\sum_{k=1}^{L} |U_{ik}^{(1)} - U_{ik}^{(r)}| \right)$$

If the distance D is 0 it means that the user final opinion did not depend on the different sequence of messages.

We see the indication of a breaking of replica symmetry, revealed by the average distribution of D_i as we can see in Figure 7. One can see that for a larger number of users with respect to factors (as happens in popular social media), the differences among the average final factors for the same user is almost never zero. This implies that the recommendation system does actually determine the final structure of user factors, in a way that is strongly dependent on the particularities of the interactions.
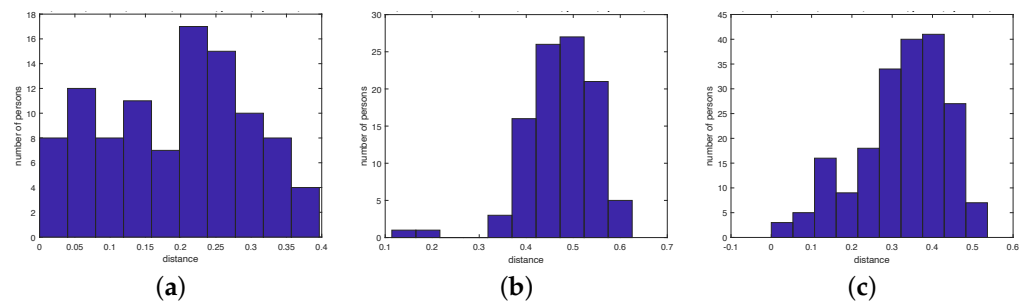
Figure 7. Histograms of replica simulations made starting from same set of random users, $R = 10$ repetitions and $M_0 = 200$, $\tau = 0.4$, $f = 0.4$, $\varepsilon = 0.3$, $M = 20{,}000$. For each of these simulations, we varied number of users and number of internal factors (opinions). (**a**) $N = 100$, $L = 40$; (**b**) $N = 100$, $L = 20$; (**c**) $N = 200$, $L = 40$. On x-axes, there is distance between users in replica 1 and same users in other other replicas.

4. Conclusions

We studied the influence of a recommendation system on the formation of originally nonexisting communities in a simulated social media system with collaborative filtering.

In our model, people are modeled as simple linear perceptrons, and are characterized by a set of factors, i.e., preferences or tastes. A message is similarly formed by a set of components, and the opinion about a message is given by the match between factors and components. The message emitted by an individual is formed by a subset of his/her weights. The recommendation system delivers messages only to those people that are expected to express a positive judgment according to the database of past opinions. People then have a certain probability of evolving their weights accordingly.

We always start by assigning random weights to people, so there is no initial community structure. When the recommendation system is initialized with a limited number of random posts, apparent communities may appear due to fluctuations in the sampling. When the recommendation system is asked to deliver messages only to people that are expected to express a highly positive opinion, those messages are sent only to those belonging to the fake community, which is therefore strengthened in the database. However, in this case the final state of the community structure reflects the real user overlap.

However, as already suggested for different systems [31] the clustering effect, i.e., bubble formation, becomes real when people are let free to evolve their factors in the "direction" of the incoming messages. In this case, real communities form and stabilize due to the interplay between the initial fluctuations and the recommendation system, and therefore real isolation bubbles are formed.

The community structure is strongly dependent on the sequences of messages, and does not depend uniquely on the initial distribution of user factors, i.e., if one could repeat the experience with limited changes (in this case, the order of users selected to send messages), as in the movie "sliding doors", one could get a completely different structure of final user "tastes".

Once a user is connected to other users via the recommendation system, their factors evolve, leading to the formation of communities that were not present in the original distribution of factors. The velocity of the bubble formation depends mainly on the value of the factor evolution parameter that models the memory, while the numbers of community increases when the selection threshold of the recommendation system decreases.

Recommendation systems do indeed have an influence on our lives. Probably, the fear for the "inevitable" formation of filter bubbles or echo chambers is exaggerated or not universally applicable, but there is indeed an influence at least in the formation of communities.

The result is that the very first interactions in a social network with a recommendation system may determine an isolation bubble not univocally determined by the initial opinion

of the users, and even promote the formation of communities only based on the random sampling of messages.

The consequence is that the first "posts" on a social group determine the final community structure.

The current model is extremely simplified, which makes the interpretation of results simpler and clearer, with the risks of losing information related to the interaction of a much larger number of users or users with many more internal parameters. Moreover, the only possible evolution in this model is through the exchange of messages on social networks, ignoring the possibility of external input. We believe that the reported effects should not vary with a more realistic number of users or parameters.

Anyway, further simulations and an experimental verification of the effect seen in the simulation might cast more light on the possibility of the real occurrence of these events.

References

1. Sindermann, C.; Cooper, A.; Montag, C. A short review on susceptibility to falling for fake political news. *Curr. Opin. Psychol.* **2020**, *36*, 44–48. [CrossRef]
2. Xu, C.; Li, J.; Abdelzaher, T.F.; Ji, H.; Szymanski, B.K.; Dellaverson, J. The Paradox of Information Access: On Modeling Social-Media-Induced Polarization. *arXiv* **2020**, arXiv:2004.01106.
3. Patel, B.; Desai, P.; Panchal, U. Methods of recommender system: A review. In Proceedings of the 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 17–18 March 2017. [CrossRef]
4. Berman, R.; Katona, Z. Curation Algorithms and Filter Bubbles in Social Networks. *Mark. Sci.* **2020**, *39*, 296–316. [CrossRef]
5. Kahneman, D. *Thinking, Fast and Slow*; Farrar, Straus and Giroux: New York, NY, USA, 2011.
6. Pariser, E. *The Filter Bubble: What the Internet Is Hiding from You*; Penguin Press: New York, NY, USA, 2011.
7. Bingbing, T.; Tianlong, G.; Yan, L. Research on Consumers' Response to Personalized Recommendation Avoidance in B2C E-business under Filter Bubble Phenomenon. In Proceedings of the 2019 International Conference on E-Business and E-Commerce Engineering, Bali, Indonesia, 21–23 December 2019. Available online: https://dl.acm.org/doi/abs/10.1145/3385061.3385068 (accessed on 16 November 2021).
8. Cinelli, M.; Morales, G.D.F.; Galeazzi, A.; Quattrociocchi, W.; Starnini, M. The echo chamber effect on social media. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2023301118. [CrossRef]
9. Sasahara, K.; Chen, W.; Peng, H.; Ciampaglia, G.L.; Flammini, A.; Menczer, F. Social influence and unfollowing accelerate the emergence of echo chambers. *J. Comput. Soc. Sci.* **2020**, *4*, 381–402. [CrossRef]
10. O'Hara, K.; Stevens, D. Echo Chambers and Online Radicalism: Assessing the Internet's Complicity in Violent Extremism. *Policy Internet* **2015**, *7*, 401–422. [CrossRef]
11. Bruns, A. Filter bubble. *Internet Policy Rev.* **2019**, *8*, 4. [CrossRef]
12. Bozdag, E.; van den Hoven, J. Breaking the filter bubble: democracy and design. *Ethics Inf. Technol.* **2015**, *17*, 249–265. [CrossRef]
13. Irakleous, G. Algorithmic Culture and Filter Bubble: The Case of YouTube's Recommendation System. Bachelor's Thesis, Ktisis Cyprus University of Technology, Limassol, Cyprus, 2020. Available online: https://ktisis.cut.ac.cy/handle/10488/18479 (accessed on 16 November 2021).
14. Lunardi, G.M.; Machado, G.M.; Maran, V.; de Oliveira, J.P.M. A metric for Filter Bubble measurement in recommender algorithms considering the news domain. *Appl. Soft Comput.* **2020**, *97*, 106771. [CrossRef]
15. Valdez, A.C.; Ziefle, M. Human Factors in the Age of Algorithms. Understanding the Human-in-the-loop Using Agent-Based Modeling. In *Social Computing and Social Media. Technologies and Analytics*; Springer International Publishing: New York, NY, USA, 2018; pp. 357–371. [CrossRef]
16. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386–408. [CrossRef]
17. Bagnoli, F.; Berrones, A.; Franci, F. De gustibus disputandum (forecasting opinions by knowledge networks). *Phys. A Stat. Mech. Its Appl.* **2004**, *332*, 509–518. [CrossRef]
18. Minsky, M.; Papert, S.A. *Perceptrons: An introduction to Computational Geometry*; MIT Press: Cambridge, MA, USA, 2017.

19. Thurstone, L.L. Multiple factor analysis. *Psychol. Rev.* **1931**, *38*, 406–427. [CrossRef]
20. Hotelling, H. Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol.* **1933**, *24*, 417–441. [CrossRef]
21. Bandalos, D.L. *Measurement Theory and Applications for the Social Sciences*; Guilford Publications: New York, NY, USA, 2018.
22. Thurstone, L.L. The vectors of mind. *Psychol. Rev.* **1934**, *41*, 1–32. [CrossRef]
23. Maslov, S.; Zhang, Y.C. Extracting Hidden Information from Knowledge Networks. *Phys. Rev. Lett.* **2001**, *87*, 248701. [CrossRef] [PubMed]
24. Blattner, M.; Zhang, Y.C.; Maslov, S. Exploring an opinion network for taste prediction: An empirical study. *Phys. A Stat. Mech. Its Appl.* **2007**, *373*, 753–758. [CrossRef]
25. Deffuant, G.; Neau, D.; Amblard, F.; Weisbuch, G. Mixing beliefs among interacting agents. *Adv. Complex Syst.* **2000**, *03*, 87–98. [CrossRef]
26. Hegselmann, R.; Krause, U. Opinion dynamics and bounded confidence models, analysis, and simulation. *J. Artif. Soc. Soc. Simul.* **2002**, *5*, 3.
27. Sznajd-Weron, K.; Sznajd, J. Opinion Evolution in Closed Community. *Int. J. Mod. Phys. C* **2000**, *11*, 1157–1165. [CrossRef]
28. Maia, H.; Ferreira, S.; Martins, M. Adaptive network approach for emergence of societal bubbles. *Phys. A Stat. Mech. Its Appl.* **2021**, *572*, 125588. [CrossRef]
29. Geschke, D.; Lorenz, J.; Holtz, P. The triple-filter bubble: Using agent-based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *Br. J. Soc. Psychol.* **2018**, *58*, 129–149. [CrossRef] [PubMed]
30. Gottron, T.; Schwagereit, F. The Impact of the Filter Bubble—A Simulation Based Framework for Measuring Personalisation Macro Effects in Online Communities. *arXiv* **2016**, arXiv:cs.SI/1612.06551.
31. Vicente, R.; Martins, A.C.R.; Caticha, N. Opinion dynamics of learning agents: Does seeking consensus lead to disagreement? *J. Stat. Mech. Theory Exp.* **2009**, *2009*, P03015. [CrossRef]