Feed for good?
On regulating social media platforms*

Miguel Risco[†] Manuel Lleonart-Anguix[‡]

July 9, 2023

Abstract

Social media platforms govern the exchange of information between users by providing personalized feeds. This paper shows that the pursuit of engagement maximization, driven by monetary incentives, results in low-quality communication and the proliferation of echo chambers. A monopolistic platform disregards social learning and curates feeds that primarily consist of content from like-minded individuals. We study the consequences on learning and welfare resulting from transitioning to this algorithm from the previously employed chronological feed. We show that the platform could create value by using its privileged information to design algorithms that balance learning and engagement, maximizing users' welfare. However, incentivizing a monopolist to embrace such an approach presents challenges. To address this, we propose interoperability as a measure to overcome network effects in platform competition, level the playing field, and prompt platforms to adopt the socially optimal algorithm.

Keywords: social learning, personalized feed, platform competition, network effects, interoperability

JEL Codes: D43, D85, L15, L86.

*Risco acknowledges financial support by the German Research Foundation (DFG) through CRC TR 224 (Project B05). We thank Sven Rady, Pau Milán, Francesc Dilmé, Alexander Frug, Carl-Christian Groh, Philipp Strack, Manuel Mueller-Frank, David Jiménez-Gómez, Fernando Payro, Jordi Caballé, Raquel Lorenzo and Diego Fica and participants at the Workshop on The Economics of Digitalization, the EDP Jamboree, the Conference on Economic Design, various CRC TR 224 Young Research Workshops, the ENTER Jamboree, the BSE Jamboree, as well as workshops at the University of Bonn, Pompeu Fabra University, and Autonomous University of Barcelona for their helpful comments and discussions.

[†]Bonn Graduate School of Economics, University of Bonn. risco@uni-bonn.de

[‡]IDEA (Autonomous University of Barcelona) and BSE. manuel.lleonart@bse.eu

1 Introduction

More than a third of Americans get their news from Facebook[1], which has 2.9 billion users worldwide. Another platform owned by Meta, Instagram, has 1.1 billion users. While both platforms were once considered harmless places for communication and content sharing, this is no longer the case. They have been criticized for causing polarization and spreading misinformation, promoting echo chambers, and fueling hate speech[2]. The 2016 US presidential election was a significant example (Solon, 2016), as Facebook was accused of failing to combat fake news. However, the biggest issue is not the leniency of the platforms, but the impact of personalized content. The *News Feed* is a customized scroll of friends' content and news stories that appears on most social media platforms. Since 2018, it is no longer chronological, but instead, a proprietary algorithm controls what appears on the feed. The algorithm considers factors such as the users' friends, joined groups, liked pages, targeted advertisers, and popular stories to provide a personalized feed[3]. Since platforms' revenues come from advertising, their primary goal is to maximize engagement, which may not align with promoting informative communication. When the feed was chronological, observing a portion of one's neighbors could be representative of the whole community or even society. However, this is no longer the case, and it seems likely that under personalized feeds, agents will be biased and make incorrect extrapolations. The following quote explains this tension. Lauer (2021): "if Facebook employed a business model focused on efficiently providing accurate information and diverse news, rather than addicting users to highly engaging content within an echo chamber, the algorithmic outcomes would be very different".

Two main tendencies have emerged when proposing policies to alleviate this situation: intervening directly on the information that is passed on and intervening on the platform structure. Information-targeting policies include censoring, fact-checking, nudging, or providing platform-generated content. Structural interventions could consist of capping the depth or breadth of the network, shutting down certain communities (deplatforming), or regulating the level of homophily that the algorithm can induce. Our work follows a market structure approach, examining the incentives of an engagement-maximizing monopolist platform when providing users with a personalized feed. We find that the monopolist optimally provides a feed in which the messages of the most similar neighbors are shown, often making social learning less informative than the traditional random feed where users observe friends' messages in chronological order[4]. This is a

1 See Gottfried and Shearer (2019).

2 See Silverman (2016) or Allcott and Gentzkow (2017).

3 For an in-depth investigation of Facebook's algorithm and its negative effects, see Horwitz et al. (2021) based on a review of internal documents.

4 Some platforms, as Facebook or Instagram, provide users with a personalized feed by default and it is not possible to fully implement a chronological feed on one's device, while Twitter allows the user

socially sub-optimal outcome, so we investigate whether competition can incentivize platforms to adopt the socially optimal algorithm (which we define as simply the one that maximizes users' sum of utilities). However, due to network effects, in equilibrium, platforms choose the same algorithm and divide the user base among them. To address this issue, we propose interoperability as a solution. Interoperability allows for total information exchange between platforms, removing network effects. With interoperability, competition disciplines platforms, forcing them to use the socially optimal algorithm in equilibrium.

It is worth noting that most successful platforms, such as Facebook, Instagram, TikTok, and Twitter, are enormous in size and monopolies in their respective fields[5]. Although these platforms can be broadly described as social media platforms that enable public posting and private communication, they differ in core functionality. Each site dominates a specific field: photography (Instagram), videos (TikTok), reciprocal communication with friends (Facebook), and micro-blogging (Twitter). When a new platform emerges, it either finds a niche and quickly becomes a monopolist or must compete with the incumbent. In the latter case, the entrant either has significantly better quality and can overcome existing network effects (e.g. Facebook vs. MySpace) or the incumbent is protected by network effects (e.g. Twitter vs. Mastodon). Even though a large platform might be providing a suboptimal service through its personalized feed, its size makes it up against a competing platform with a better service but smaller user base[6]. Interoperability can put an end to this phenomenon by allowing users to choose the platform that provides them with the best service, regardless of its size. With interoperability, competing platforms would be forced to implement the socially optimal algorithm, as otherwise, users would leave.

We highlight three main contributions of this paper: first, we build a model of *communication and learning through personalized feed*, where an strategic engagement-maximizing platform chooses an algorithm and users post messages. We assume that an individual (she) joins a social media platform for two main reasons. First, to engage in communication with peers about some underlying topic. This social activity generates utility through two channels: expressing one's own views (in the sense of being loyal to own innate opinions; *sincerity*), and conforming with the rest (in the sense of matching the opinions that neighbors have shared; *conformity*[7]). Second, to learn about some state

to choose directly in the timeline screen.

5 Regarding monopoly structures in the social media platform market, a quote from the Bundeskartellamt (the German competition protection authority) in its case against Facebook (B6-22/16, "Facebook", p. 6) states: "The facts that competitors are exiting the market and there is a downward trend in the user-based market shares of remaining competitors indicate a market tipping process that will result in Facebook becoming a monopolist". (Franck and Peitz, 2022).

6 The outside option of leaving social media platforms is a difficult decision for an individual due to the fear of missing out (Przybylski et al., 2013)

of the world and make the best decision outside the platform. The Covid-19 vaccine could serve as an example: even though vaccine effectiveness is a purely scientific matter, there has been a considerable public debate about it[8], with individuals sharing their views on social media platforms for both sincerity and conformity, but also for learning purposes. Learning is needed to make (the best possible) decision on getting vaccinated or not. Maximizing revenues means maximizing user engagement[9], which is a function of the utility users derive from interacting on the platform. The idea is that a pleased user will return to the platform. In contrast, learning is seen as a long-term reward and does not affect engagement. Users communicate by sending messages about the topic of interest and learn through reading the messages that appear in their personalized feed. The feed is a subset of neighbors' messages the platform chooses. We assume that users' private information is platform's knowledge, so it can design the feed conditional on it. The degree of machine learning techniques sophistication and the amount of data available justify this assumption.[10] We find that in equilibrium, a revenue-maximizing monopolist platform chooses an algorithm that shows to each user the messages of her most similar neighbors—we call this algorithm "closest" algorithm. In turn, users report truthfully their innate opinion. We also show that, surprisingly, this algorithm does not always harm social learning when users are sophisticated. Still, the "user optimal" algorithm is socially preferred, so we wonder how this outcome could be implemented. This leads to the second contribution of the paper.

Based on the platform-users game introduced above, we build a simple extension to study competition. Platforms simultaneously decide on which algorithm to implement, and then users choose which platform to join. Afterwards, the game of *communication and learning through personalized feed* takes place. We find that the existence of network effects permits platforms to keep playing the "closest" algorithm in equilibrium. Even if a platform implements the socially optimal algorithm, a larger user base could make the "closest" algorithm implemented by the other platform preferable for a given user. Thus, competition is not enough to implement the socially optimal algorithm in equilibrium. Intervention is then justified, and we look for a regulation policy that might work. This

7 Conformity is a driving-force in social media behavior (Mosleh et al., 2021). It is defined as the act of matching attitudes, beliefs and behaviors to group norms (Cialdini and Goldstein, 2004). Here we treat conformity as a behavioral bias included at the outset, but it has been widely found as a product of rational models. See Bernheim (1994) for a theory of conformity and Chamley (2004) for an overview.

8 The acceptance of the Covid-19 vaccine in US and UK declined an average of 6 percentage points due to misinformation (Loomba et al., 2021).

9 An ad-revenue-maximizing platform as in Mueller-Frank et al. (2022) maximizes revenues by maximizing the amount of users that observe advertisements.

10 Facebook's FBLearner Flow, a machine learning platform, is able to predict user behavior through the use of personal information collected within the platform. See Biddle (2018) for a news piece on it. The early paper Kosinski et al. (2013) already showed that less sophisticated techniques could predict a wide range of personal attributes by just using data on "likes".

is our third contribution. We claim that the enforcement of interoperability removes network effects and forces competing platforms to implement the socially optimal algorithm. Interoperability allows platforms to connect, so that users from different platforms can be neighbors. Hence, users will choose the platform that provides with the best service, disregarding how many users it has. Consequently, any platform must implement the socially optimal algorithm in equilibrium, as otherwise it would be empty.

The rest of the paper is organized as follows. After the literature review, Section 2 introduces our basic environment for a monopoly platform and n users game. Section 3 analyzes the equilibrium of that game and Section 4 characterizes the main algorithms appearing in the model. Section 5 analyzes platform competition and interoperability and Section 6 concludes.

1.1 Related literature

This paper is related to two areas of literature. The first area studies the impact of revenue-maximizing platforms on social learning. This is a growing field, and we highlight two papers for their similarities to our work. In a model where agents decide whether or not to pass on (mis)information, Acemoglu et al. (2021) studies the algorithm choice of the platform, which maximizes engagement. They show that when the platform has the ability to shape the network (which is equivalent to choosing personalized feeds, but without constraints on pre-existing neighborhoods), it will design algorithms that create more homophilic communication patterns, forming echo chambers. Mueller-Frank et al. (2022) build a model of network communication and advertising where the platform controls the flow of information. In equilibrium, the platform may manipulate or even suppress information to increase revenue, even though this ultimately decreases social welfare. Additional research on media platforms providing distorted content for economic reasons can be found in Reuter and Zitzewitz (2006), Ellman and Germano (2009), Abreu and Jeon (2019), and Kranton and McAdams (2020). The topic of homophilic communities and echo chambers is discussed in Sunstein (2017). Hu et al. (2021) shows that rational, inattentive users prefer to learn from like-minded neighbors, while Törnberg (2018) shows that echo chambers harm social welfare by increasing the spread of misinformation.

Not just the mentioned literature, but also empirical work (cf. Sagioglou and Greitemeyer (2014) or Levy (2021)) reveals the need for further intervention or regulation on social media platforms. This topic constitutes the second strand to which our paper is closely related. Franck and Peitz (2022) puts context to social media platform competition, claiming that market power (mainly represented by the network effects) leads to suboptimal outcomes for society. The current mechanics suggest that it may not be the platform with the best offer that dominates the market. Popiel (2020) and Evens et al. (2020) assert that regulations to manage digital platform markets in the US and EU, respectively, are inadequate in addressing their negative effects. In response to this need, there has been a surge of recent papers examining interventions. Regarding structural interventions, Jackson et al. (2022) examines how limiting the breadth and/or depth of a social network improves message accuracy. The work of Benzell and Collis (2022) aligns with our own, as they analyze the optimal strategy of a monopolist social media platform and evaluate the impact of taxation and regulatory policies on both platform profits and social welfare. However, in their model, the platform chooses net revenue per user rather than shaping communication among users. The authors apply their model to Facebook and find that a successful regulatory intervention to achieve perfect competition would increase social welfare by 4.8%, which supports our theoretical findings. Finally, Agarwal et al. (2022) provides empirical evidence of the negative consequences of deplatforming (shutting down a community on a platform), mainly due to migration effects, which supports our call for globally applicable regulations.

There is a plethora of recent empirical contributions regarding informational interventions: Habib et al. (2019), Hwang and Lee (2021) or Mudambi and Viswanathan (2022). Mostagir and Siderius (2023b) models community formation and shows that the effect of interventions is non-monotonic over time. Additionally, there is another important aspect to consider when analyzing informational policies: Mostagir and Siderius (2022) demonstrates that cognitive sophistication matters when faced with misinformation, and Mostagir and Siderius (2023a) finds that different populations (Bayesian and DeGrootians) react differently to certain interventions. While some papers, such as Mostagir and Siderius (2023a), include cases where sophisticated users are outperformed by their naive counterparts, Pennycook and Rand (2019) and Pennycook and Rand (2021) show that higher cognitive ability is associated with better ability to discern fake content. In our model, the results hold for both Bayesian and DeGrootian users, but the sophisticated users always learn better. Finally, we also relate to the literature on learning in networks, for both naive and sophisticated users: DeMarzo et al. (2003), Acemoglu and Ozdaglar (2011), Jadbabaie et al. (2012), Molavi et al. (2018) or Mueller-Frank and Neri (2021).

2 Model

We first outline the model. There is an underlying state of the world denoted by $\theta \in \mathbb{R}$, and every user receives a private signal about it. Users join a social media platform to communicate through messages, deriving utility through two streams. Within-the-platform utility depends on the posted messages and the users' original opinions about θ. The larger the within-the-platform utility, the more engaged a user is. Outside-the-platform utility, referred to as learning, is given by a quadratic loss function accounting for the deviation of the user's action from θ. Such an action is based on both private signals and the messages learnt within the platform. We could interpret within-the-platform utility as *media user* utility, representing the immediate payoff a myopic consumer obtains while spending time in the platform. Total utility, on the other hand, corresponds to a sort of *citizen* utility, capturing the overall payoff a conscious agent derives from her entire experience on the platform. Considering these aspects of user utility, the platform selects a subset of neighbors whose messages appear in each user's personalized feed by leveraging information on users' similarity. The platform's revenue is given by total engagement, so its primary concern is maximizing users' within-the-platform utility.

Now, let us describe the model in detail. There is a set, N, of n users that join a social media platform described by the (un)directed graph $\mathcal{G}$. Each node in the graph represents a user in the platform, and two nodes are linked if and only if they are friends on the platform. The set of user i's friends—her neighborhood—is called N_i. Each user receives a private signal about θ, $\theta_i \in \mathbb{R}$. Conditional on θ, signals $\{\theta_1, ..., \theta_n\}$ are jointly normal and their structure is given by:

$$\begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_n \end{pmatrix} \sim \mathcal{N}\left(\boldsymbol{\theta}, \boldsymbol{\Sigma}\right),$$

where $\boldsymbol{\theta} = (\theta, ..., \theta)$ and $\boldsymbol{\Sigma}$ is an $n \times n$ symmetric and positive definite matrix where $\Sigma_{ii} = \sigma^2$ for every i and $\Sigma_{ij} = \text{Cov}(\theta_i, \theta_j)$ for every i, j. We denote by ρ_{ij} the correlation between θ_i and θ_j, so that $\rho_{ij} = \frac{\Sigma_{ij}}{\sigma^2}$ for every i, j. We assume improper priors[11] on θ. The random variable θ_i is user i's private opinion on θ, which is based on inherent personal characteristics but also on information collected privately[12]. Users know their private signals, the distribution of all signals, the covariance matrix $\boldsymbol{\Sigma}$, and the distribution of the state of the world. Thus, user i's posterior distributions are $\theta_j|\theta_i \sim \mathcal{N}\left(\theta_i, \sigma^2(1-\rho_{ij})\right)$ for all $j \in N$ and $\theta|\theta_i \sim \mathcal{N}(\theta_i, \sigma^2)$.

Regarding the communication phase, each user i simultaneously posts a message $m_i \in \mathbb{R}$. Then, she observes her personalized feed. This subset is strategically chosen by the platform through an algorithm that will be defined later. In reality, social media platforms order (rank) friends' posts and users observe as many as their scrolling time allows them to. For the purpose of this analysis, we assume that all users spend a fixed amount of time reading messages, i.e., the number of observed messages is exogenously fixed, and that the amount of available friends' posts exceeds such a number. For simplicity, we work under the assumption that every user observes the same number of messages, but this is without loss of generality. The key assumption here is exogeneity: neither the users nor the platform are able to adjust the number of messages that will be read. Although

11 For a discussion of improper priors, see Hartigan (1983).

12 The signal θ_i can be interpreted as the information the user has about the state of the world prior to her entry on the social platform. As information sources, as well as ideology, might be similar, different users' private information might be correlated. This is captured by the matrix $\boldsymbol{\Sigma}$.

user behavior on social media platforms is highly diverse, existing data support these assumptions[13].

Users derive utility after communication takes place, following two payoff streams; (i) *sincerity*: agents are punished for deviating from their own signals, and (ii) *conformity*: matching others' opinions is rewarded. This utility is called within-the-platform utility and is given by

$$u_i(m_i; m_{-i}, \mathcal{S}_i) = -\beta \overbrace{(\theta_i - m_i)^2}^{\text{Sincerity}} - (1-\beta) \overbrace{\sum_{j \in \mathcal{S}_i} \frac{(m_i - m_j)^2}{k}}^{\text{Conformity}}, \tag{1}$$

where $\beta \in (0,1)$ represents the weight each of the streams receives and $\mathcal{S}_i$ denotes user i's personalized feed. We assume the cardinality of all $\{\mathcal{S}_i\}_{i=1}^n$ to be exogenously set to $k \in \mathbb{N}$. Within-the-platform utility is not the only source of utility for users, as they are also concerned about taking an action that matches the state of the world. Platform communication allows them to learn about θ and optimize the decision making proccess. Deciding on the action conditional on the messages learnt is what we call learning. Total utility is the weighted average of within-the-platform utility and learning:

$$U_i(m_i, m_{-i}, a_i, \mathcal{S}_i) = \lambda \left[-\beta(\theta_i - m_i)^2 - (1-\beta) \sum_{j \in \mathcal{S}_i} \frac{(m_i - m_j)^2}{k} \right] - (1-\lambda) \overbrace{(a_i - \theta)^2}^{\text{Learning}}. \tag{2}$$

The optimal action a_i^* is the best guess of θ conditional on the observed messages, and $\lambda \in (0,1)$ weights the relative importance of within-the-platform and learning utilities. Thus, user i chooses a message m_i and, after learning messages $\{m_j\}_{j \in \mathcal{S}_i}$, she chooses an action a_i to maximize her (expected) utility U_i.

Next, we introduce the **platform** as a strategic agent of the game. It knows $\mathcal{G}$ and $\boldsymbol{\Sigma}$, but not θ or $\{\theta_i\}_{i=1}^n$. For each user i it chooses the set of neighbors that she will observe messages from. As explained above, this set is called personalized feed and denoted by $\mathcal{S}_i$, has cardinality k (so it is a subset of N_i), and $\mathcal{G}$ is such that $k \leq \min_{i \in N}\{|N_i|\}$ for all i. We define the platform's *algorithm* as the mapping that provides for each user a feed:

$$\begin{array}{rccc} \mathcal{F}: & N & \longrightarrow & \prod_{i=1}^n N_i \\ & (1,2,...,n) & \mapsto & (\mathcal{S}_1, ..., \mathcal{S}_n). \end{array}$$

Engagement is defined as an increasing non-negative function of the sum of every user's within-the-platform utility. As explained above, the intuition hinges on the fact that the happier a user is in the platform, the higher the probability of her coming back

13 According to the statistics portal *Statista*, an average Facebook user has 335 friends and spends 35 minutes on the platform, reading between 10 and 50 posts. However, at least 300 stories are produced in her network. Finally, on average, users tend to consistently allocate a similar amount of screen time to social media platforms.

in the subsequent periods. For simplicity, we assume that such a function is the identity and as there are no costs in the model, profits are directly given by engagement:

$$\Pi_p(m, \mathcal{F}) = \sum_{i=1}^{n} u_i(m_i, m_{-i}, \mathcal{S}_i).$$

The game of *communicating and learning through personalized feed* described above is played by the platform and the users, and it consists of the following sequence of events:

1. The platform chooses an algorithm $\mathcal{F}$ and commits to it.
2. Each user observes her private signal θ_i.
3. Each user i sends a message $m_i \in \mathbb{R}$.
4. Each user i observes the messages in her feed $\mathcal{S}_i$ and chooses an action a_i.
5. The state of the world is revealed and payoffs are realized.

3 Equilibrium

Here we describe and analyze the equilibrium of the game. The equilibrium concept is Bayes-Nash linear equilibrium. The platform's strategy consists of the choice of an algorithm $\mathcal{F}$. In turn, each user decides on a pair (m_i, a_i) consisting of a message and an action that maximize her expected payoff given $\mathcal{F}$. Note that at the time of choosing the action, the user has learnt the messages posted in her personalized feed. An equilibrium is given by pairs of messages and actions for the users and an algorithm for the platform. Note that the platform chooses an algorithm that maximizes its benefits given the induced equilibrium strategies of the users in the subsequent subgame. We find that in equilibrium, users report truthfully their private signal and the optimal algorithm for the platform is one that shows, for each user, the messages of those neighbors who feature the highest correlation with her (in other words, the most similar, or the *closest* friends). We refer to such an algorithm as the "closest" algorithm $\mathcal{C}$.

Before formally deriving the equilibrium of the game, let us show two auxiliary results. First, we prove that when users report truthfully their types, it is equivalent for the platform to maximize engagement and to maximize each user's within-the-platform utility separately. I.e., there are no inter-dependencies across feeds. This result also implies that an algorithm that maximizes each user's individual utility is precisely the utilitarian optimal algorithm that a social planner would implement.

Lemma 3.1. *If $m_i = \theta_i$ for all $i \in N$, then*[14]

$$\underset{(\mathcal{S}_1,\ldots,\mathcal{S}_n)}{\operatorname{argmax}} \left\{ \sum_{i=1}^{n} u_i \right\} = \left(\underset{\mathcal{S}_1 \subseteq N_1}{\operatorname{argmax}}\{u_1\}, \ldots, \underset{\mathcal{S}_n \subseteq N_n}{\operatorname{argmax}}\{u_n\} \right).$$

14 We assume, without loss of generality, that the feeds $\{\mathcal{S}_i\}_{i=1}^{n}$ that maximize engagement are unique.

Proof. By definition of within-the-platform utility, if types are reported truthfully the only feed that affects user i is $\mathcal{S}_i$. Then,

$$\max_{(\mathcal{S}_1,\ldots,\mathcal{S}_n)}\left\{\sum_{i=1}^{n}\mathbb{E}_p[u_i]\right\} = \max_{(\mathcal{S}_1,\ldots,\mathcal{S}_n)}\left\{\sum_{i=1}^{n}\mathbb{E}_p\left[-\beta(\theta_i-m_i)^2-(1-\beta)\sum_{j\in\mathcal{S}_i}\frac{(m_i-m_j)^2}{k}\right]\right\} =$$
$$= \max_{(\mathcal{S}_1,\ldots,\mathcal{S}_n)}\left\{-\sum_{i=1}^{n}\left(\frac{1-\beta}{k}\sum_{j\in\mathcal{S}_i}\mathbb{E}_p[(\theta_i-\theta_j)^2]\right)\right\} = \sum_{i=1}^{n}\left(\max_{(\mathcal{S}_1,\ldots,\mathcal{S}_n)}\left\{-\frac{1-\beta}{k}\sum_{j\in\mathcal{S}_i}\mathbb{E}_p[(\theta_i-\theta_j)^2]\right\}\right) =$$
$$= \sum_{i=1}^{n}\left(\max_{\mathcal{S}_i\subseteq N_i}\{\mathbb{E}_p[u_i]\}\right).$$

□

Next, we derive the distribution of θ conditional on the messages observed in the personalized feed if there is truthful reporting, i.e., $\{\theta_j\}_{j\in\mathcal{S}_i}$.

Lemma 3.2. *The posterior distribution of θ conditional on $\{\theta_j\}_{j\in\mathcal{S}_i}$ is given by*

$$\theta|\{\theta_j\}_{j\in\mathcal{S}_i}\sim\mathcal{N}\left(\frac{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\boldsymbol{\theta}_{\mathcal{S}_i}}{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}},\frac{1}{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}}\right),$$

where $\mathbb{1}$ is a n-vector of ones, $\boldsymbol{\Sigma}_{\mathcal{S}_i}$ is the submatrix of $\boldsymbol{\Sigma}$ induced by $\mathcal{S}_i$ and $\boldsymbol{\theta}_{\mathcal{S}_i} = (\theta_{j_1}\ldots\theta_{j_n})$, with $j_r\in\mathcal{S}_i$.

Proof. Let us assume, for simplicity, that the signals user i observes in her personalized feed $\mathcal{S}_i$ are $\{\theta_1,\ldots,\theta_k\}$. We know that $(\theta_1\ldots\theta_k)\sim\mathcal{N}(\boldsymbol{\theta},\boldsymbol{\Sigma}_{\mathcal{S}_i})$. Now, the posterior distribution of θ conditional on $\{\theta_j\}_{\mathcal{S}_i}$ is characterized by its pdf:

$$g(\theta|\{\theta_j\}_{\mathcal{S}_i}) = (2\pi\det(\boldsymbol{\Sigma}_{\mathcal{S}_i}))^{-1/2}\exp\left[-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}_{\mathcal{S}_i})^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}(\boldsymbol{\theta}-\boldsymbol{\theta}_{\mathcal{S}_i})\right] =$$
$$= (2\pi\det(\boldsymbol{\Sigma}_{\mathcal{S}_i}))^{-1/2}\exp\left[-\frac{1}{2}\left(\theta^2\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}-2\theta\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\boldsymbol{\theta}_{\mathcal{S}_i}+\boldsymbol{\theta}_{\mathcal{S}_i}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\boldsymbol{\theta}_{\mathcal{S}_i}\right)\right].$$

Treated as the pdf of the posterior distribution of θ, the above expression is equivalent to

$$g(\theta|\{\theta_j\}_{\mathcal{S}_i}) = \sqrt{\frac{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}\mathbb{1}}{2\pi}}\exp\left[-\frac{1}{2}\left(\theta^2\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}-2\theta\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\boldsymbol{\theta}_{\mathcal{S}_i}+\frac{(\boldsymbol{\theta}_{\mathcal{S}_i}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1})^2}{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}}\right)\right] =$$
$$= \sqrt{\frac{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}\mathbb{1}}{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{\theta-\frac{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\boldsymbol{\theta}_{\mathcal{S}_i}}{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}}}{\sqrt{\frac{1}{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}}}}\right)^2\right].$$

Thus,

$$\theta|\{\theta_j\}_{j\in\mathcal{S}_i}\sim\mathcal{N}\left(\frac{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\boldsymbol{\theta}_{\mathcal{S}_i}}{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}},\frac{1}{\mathbb{1}^t\boldsymbol{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}}\right)$$

as we wanted to show. □

Next, we derive the equilibrium of the game.

Proposition 3.3. *The unique Bayesian Nash equilibrium in linear strategies of the game is given by users playing* $\left(m_i^* = \theta_i,\ a_i^* = \frac{\mathbb{1}\Sigma_{\mathcal{S}_i}^{-1}\boldsymbol{\theta}_{\mathcal{S}_i}^t}{\mathbb{1}\Sigma_{\mathcal{S}_i}^{-1}\mathbb{1}^t}\right)$ *and the platform playing* $\mathcal{C}$*, the "closest" algorithm.*

Proof. We analyze first the optimal message for the user[15]. Given user i's type θ_i, the algorithm $\mathcal{F}$, and the matrix $\boldsymbol{\Sigma}$, she chooses a message $m_i \in \mathbb{R}$ such that it maximizes her expected within-the-platform utility, as learning is not affected by this choice:

$$\mathbb{E}_i[u_i(m_i, m_{-i}, \mathcal{F})|\theta_i, \mathcal{F}] = -\beta(\theta_i - m_i(\theta_i))^2 - (1-\beta)\frac{1}{k}\mathbb{E}_i\left[\sum_{j\in\mathcal{S}_i(\boldsymbol{\Sigma})}(m_i(\theta_i) - m_j(\theta_j))^2|\theta_i, \mathcal{F}\right] =$$

$$= -\beta(\theta_i - m_i(\theta_i))^2 - \frac{(1-\beta)}{k}\left(k\, m_i(\theta_i)^2 + \sum_{j\in\mathcal{S}_i(\boldsymbol{\Sigma})}\mathbb{E}_i\left[m_j(\theta_j)^2|\theta_i, \mathcal{F}\right] - 2m_i(\theta_i)\sum_{j\in\mathcal{S}_i(\boldsymbol{\Sigma})}\mathbb{E}_i\left[m_j(\theta_j)|\theta_i, \mathcal{F}\right]\right).$$

The first order condition with respect to m_i yields

$$m_i = \beta\theta_i + (1-\beta)\frac{1}{k}\sum_{j\in\mathcal{S}_i(\boldsymbol{\Sigma})}\mathbb{E}_i\left[m_j(\theta_j)|\theta_i, \mathcal{F}\right]. \tag{3}$$

Assuming linear messaging strategies for all users except from i ($m_j(\theta_j) = \gamma_j\theta_j + \delta_j$, for some $\gamma_j, \delta_j \in \mathbb{R}$ and all $j \neq i$), we can work further on the expectation term from (3) for each $j \in \mathcal{S}_i(\boldsymbol{\Sigma})$ (note that given the algorithm and the covariance matrix, the user anticipates which neighbors will appear in her feed) and obtain:

$$\mathbb{E}_i\left[m_j(\theta_j)|\theta_i, \mathcal{F}, j\in\mathcal{S}_i\right] = \mathbb{E}_i\left[\gamma_j\theta_j + \delta_j|\theta_i, \mathcal{F}, j\in\mathcal{S}_i\right] =$$
$$= \gamma_j\mathbb{E}_i[\theta_j|\theta_i, \mathcal{F}, j\in\mathcal{S}_i] + \delta_j = \gamma_j\theta_i + \delta_j.$$

Plugging this into (3) yields

$$m_i = \beta\theta_i + (1-\beta)\frac{1}{k}\sum_{j\in\mathcal{S}_i(\boldsymbol{\Sigma})}(\gamma_j\theta_i + \delta_j) = \beta\theta_i + (1-\beta)(\bar{\gamma}_i\theta_i + \bar{\delta}_i),$$

where $\bar{\gamma}_i := \sum_{j\in\mathcal{S}_i}\frac{\gamma_j}{k}$ and $\bar{\delta}_i := \sum_{j\in\mathcal{S}_i}\frac{\delta_j}{k}$. Hence, user i's optimal strategy is linear: $m_i(\theta_i) = \gamma_i\theta_i + \delta_i$. This leads to a system of equations given by

$$\begin{cases} \gamma_i &= \beta + (1-\beta)\bar{\gamma}_i \\ \delta_i &= (1-\beta)\bar{\delta}_i \end{cases} \quad \forall i \in \{1, ..., n\}$$

whose unique solution is $\gamma_i = 1$ and $\delta_i = 0$ for all $i \in N$. Thus, the optimal message is $m_i^* = \theta_i$ and every user reports truthfully her type. Now, the platform chooses an

15 As the utility function is additive separable and the choice of m_i does not affect that of a_i and vice versa, we can study the optimal decisions independently.

algorithm $\mathcal{F}$ such that

$$(\mathcal{S}_1, ..., \mathcal{S}_n) = \underset{(\mathcal{S}_1,...,\mathcal{S}_n)\subseteq \prod N_i}{\text{argmax}} \left\{ -\sum_{i=1}^{n} \left(\frac{1}{k} \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} \mathbb{E}_p \left[(\theta_i - \theta_j)^2 \right] \right) \right\}.$$

But by Lemma 3.1, this is equivalent to maximizing each user's within-the-platform utility, i.e., to finding

$$\mathcal{S}_i = \underset{\mathcal{S}_i \subseteq N_i}{\text{argmax}} \left\{ - \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} \mathbb{E}_p \left[(\theta_i - \theta_j)^2 \right] \right\} =$$
$$= \underset{\mathcal{S}_i \subseteq N_i}{\text{argmax}} \left\{ - \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} (2\sigma^2 - 2\text{Cov}(\theta_i, \theta_j)) \right\} = \underset{\mathcal{S}_i \subseteq N_i}{\text{argmax}} \left\{ \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} \text{Cov}(\theta_i, \theta_j) \right\} \quad \forall\, i.$$

The algorithm that induces such feeds is precisely $\mathcal{C}$, the "closest" algorithm. The platform shows each user the messages of those neighbors that present the higher correlation to her. Thus, in equilibrium, $m_i^* = \theta_i$ for all i and $\mathcal{F} = \mathcal{C}$. Next, we calculate user i's optimal action. She maximizes $\mathbb{E}[(a_i - \theta)^2]$ conditional on the observed messages. Hence, the optimal action given $\{m_j\}_{j \in \mathcal{S}_i} = \{\theta_j\}_{j \in \mathcal{S}_i}$ is $a_i^* = \mathbb{E}_i[\theta | \{\theta_j\}_{j \in \mathcal{S}_i}] = \frac{\mathbb{1}\mathbf{\Sigma}_{\mathcal{S}_i}^{-1}\boldsymbol{\theta}_{\mathcal{S}_i}^t}{\mathbb{1}\mathbf{\Sigma}_{\mathcal{S}_i}^{-1}\mathbb{1}^t}$ by Lemma 3.2. $\square$

We dedicate the next section to understanding the consequences of implementing the "closest" algorithm on learning and total utility. However, before delving into that, we conclude this section by briefly examining how the model responds when we modify the assumptions concerning signal structure and user rationality.

3.1 Different signal structures

So far, we have assumed equal variances across users with heterogeneous correlations in our model. This assumption seems most plausible to us, as differences in opinion or ideology may exert a stronger influence than differences in learning accuracy. However, we also find it interesting to explore other signal structures for our model. In general, we can relax the restriction on variances and consider any covariance matrix. Let's start with the simpler case where variances can differ, but users are uncorrelated (this can occur, for instance, in a simplified island model). Under this specification we still observe truthful reporting in equilibrium, but the platform faces a different problem:

$$\mathcal{S}_i = \underset{\mathcal{S}_i \subseteq N_i}{\text{argmax}} \left\{ - \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} \mathbb{E}_p \left[(\theta_i - \theta_j)^2 \right] \right\} =$$
$$= \underset{\mathcal{S}_i \subseteq N_i}{\text{argmax}} \left\{ - \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} (\sigma_i^2 + \sigma_j^2 - 2\text{Cov}(\theta_i, \theta_j)) \right\} = \underset{\mathcal{S}_i \subseteq N_i}{\text{argmax}} \left\{ - \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} \sigma_j^2 \right\} \quad \forall\, i.$$

Now, the platform aims to match users with neighbors who learn most accurately. In complete contrast to the main model, the platform's interests are entirely aligned with learning.

In the general case where users have different variances and are correlated, we still observe truthful reporting in equilibrium, and the platform's problem becomes:

$$\mathcal{S}_i = \underset{\mathcal{S}_i \subseteq N_i}{\operatorname{argmax}} \left\{ - \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} \mathbb{E}_p \left[(\theta_i - \theta_j)^2 \right] \right\} =$$

$$= \underset{\mathcal{S}_i \subseteq N_i}{\operatorname{argmax}} \left\{ - \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} (\sigma_i^2 + \sigma_j^2 - 2\mathrm{Cov}(\theta_i, \theta_j)) \right\} = \underset{\mathcal{S}_i \subseteq N_i}{\operatorname{argmax}} \left\{ - \sum_{j \in \mathcal{S}_i(\mathbf{\Sigma})} (\sigma_j^2 - 2\mathrm{Cov}(\theta_i, \theta_j)) \right\} \quad \forall i.$$

The platform now aims to select users who are similar but also learn accurately. Note that similarity (covariance) carries twice the weight compared to accuracy (the neighbor's variance).

3.2 Naïve users

Naïve users are mechanical individuals who share their beliefs and update them using the DeGroot rule (DeGroot, 1974). Therefore, $m_i^* = \theta_i$ and $a_i = \frac{1}{k}\sum_{j \in \mathcal{S}_i} m_j = \frac{1}{k}\sum_{j \in \mathcal{S}_i} \theta_j$. Following the reasoning in Proposition 3.3, an engagement-maximizer platform will implement the "closest" algorithm $\mathcal{C}$ in equilibrium. However, when applied to naïve users, the "closest" algorithm always harms learning.

Proposition 3.4. *The posterior variance Var* $\left[\theta | \frac{1}{k}\sum_{j \in \mathcal{S}_i^{\mathcal{C}}} \theta_j\right]$ *given the "closest" algorithm is larger than the posterior variance induced by any other algorithm.*

Proof. Before showing that $\mathrm{Var}\left[\theta | \frac{1}{k}\sum_{j \in \mathcal{S}_i^{\mathcal{C}}} \theta_j\right] \geq \mathrm{Var}\left[\theta | \frac{1}{k}\sum_{j \in \mathcal{S}_i^{\mathcal{F}}} \theta_j\right]$ for any algorithm $\mathcal{F}$, we need to characterize the posterior variance of θ when just the average of the observed messages is learnt.

The posterior distribution of the average message conditional on θ follows $\frac{1}{k}\sum_{j=1}^{k} \theta_j | \theta \sim \mathcal{N}\left(\frac{1}{k}\sum_{j=1}^{k}\theta_j, \frac{\sigma^2}{k} + \frac{1}{k}\sum_{j \in \mathcal{S}_i} \Sigma_{ij}\right)$. As we have assumed improper priors for θ,

$$\theta | \frac{1}{k}\sum_{j=1}^{k} \theta_j \sim \mathcal{N}\left(\frac{1}{k}\sum_{j=1}^{k}\theta_j, \frac{\sigma^2}{k} + \frac{1}{k}\sum_{j \in \mathcal{S}_i} \Sigma_{ij}\right).$$

Hence, $\mathrm{Var}[\theta | \frac{1}{k}\sum_{j \in \mathcal{S}_i} \theta_j] = \frac{\sigma^2}{k} + \frac{1}{k}\sum_{j \in \mathcal{S}_i} \Sigma_{ij}$. The "closest" algorithm features the largest posterior variance among any algorithm, because by definition it chooses the neighbors whose signals feature the largest covariances with that of i. $\square$

In the case of naïve users, we can state that the "closest" algorithm harms learning. Next section explores how to incentivize platforms to move from this algorithm to the "user optimal" algorithm. The results are valid both for sophisticated an naïve learners.

4 Algorithms and learning

Under $\mathcal{C}$, each agent observes a vicinity composed of her most similar neighbors, disregarding how close their messages would be to θ in expected terms. The "closest" algorithm generates echo chambers: each user learns the messages of her like-minded neighbors. Still, the user benefits from the k messages received to learn about θ. Thus, joining the platform improves learning compared to an outside option in which she only knows θ_i. Note that the monopolist platform has no incentive to care about users' learning, as it does not affect engagement. The action a_i does not play a role when the platform chooses its optimal algorithm. This might not be desirable from a social point of view.

Let us study learning in more detail. Note that choosing such an action $a_i = \mathbb{E}_i[\theta|\{\theta_j\}_{j\in\mathcal{S}_i}]$ implies that the expected value of learning is precisely the conditional variance of θ given what the user learns by reading her personalized feed $\mathcal{S}_i$ induced by some algorithm $\mathcal{F}$, i.e., $\{\theta_j\}_{j\in\mathcal{S}_i}$:

$$\mathbb{E}\left[(a_i-\theta)^2|\{\theta_j\}_{j\in\mathcal{S}_i}\right] = \mathbb{E}\left[(\mathbb{E}_i[\theta|\{\theta_j\}_{j\in\mathcal{S}_i}]-\theta)^2|\{\theta_j\}_{j\in\mathcal{S}_i}\right] = \mathrm{Var}\left[\theta|\{\theta_j\}_{j\in\mathcal{S}_i}\right].$$

Thus, algorithms could be ranked in learning terms by comparing the conditional variance they induce. Let us define the "user optimal" algorithm ($\mathcal{U}$) as the one that maximizes each user's (expected) utility. It is the best service a platform can provide with to the user. For each user i, the algorithm induces the following feed:

$$\mathcal{S}_i^{\mathcal{U}} = \underset{\mathcal{S}_i\subseteq N_i}{\operatorname{argmax}}\left\{-\lambda(1-\beta)\frac{1}{k}\sum_{j\in\mathcal{S}_i}\mathbb{E}_p[(\theta_i-\theta_j)^2]-(1-\lambda)\,\mathbb{E}_p[(a_i-\theta)^2]\right\}.$$

Then, as $a_i = \mathbb{E}_i[\theta|\{\theta_j\}_{j\in\mathcal{S}_i^{\mathcal{U}}}]$ and the only feed that affects user i's learning is $\mathcal{S}_i$, we can write the following corollary to Lemma 3.1:

Corollary 4.1. *The "user optimal" algorithm is the utilitarian optimal algorithm:*

$$\underset{(\mathcal{S}_1^{\mathcal{U}},\ldots,\mathcal{S}_n^{\mathcal{U}})}{\operatorname{argmax}}\left\{\sum_{i=1}^{n}U_i\right\} = \left(\underset{\mathcal{S}_1^{\mathcal{U}}\subseteq N_1}{\operatorname{argmax}}\{U_1\},\ldots,\underset{\mathcal{S}_n^{\mathcal{U}}\subseteq N_n}{\operatorname{argmax}}\{U_n\}\right).$$

We highlight another algorithm, the "random" algorithm ($\mathcal{R}$), because of its (past) relevance. The "random" algorithm provides user i with a feed consisting of k messages chosen arbitrarily in N_i. It is the feed that was implemented in all platforms before personalized algorithms were introduced in the late 2010s. We refer to Appendix A for an explicit example of how the algorithms provide personalized feeds to users.

In terms of overall utility, the "user optimal" algorithm is weakly better than the rest. It is socially preferred, and we will analyze whether it is possible to incentivize platforms to implement it. However, we also devote our attention to study the social consequences of the implementation of the "closest" algorithm over the "random" algorithm. The "closest" algorithm always induces larger within-the-platform utility. In terms of learning, we could think that it works the other way around: as the "closest" algorithm creates an echo chamber for each user, learning must be harmed. Surprisingly, it is not the case in general. There are some structures for the signal technology such that learning is improved if the "closest" algorithm is implemented.

The following simple example illustrates such a case. We consider a neighborhood composed by four neighbors, and $k = 3$. The distribution of the signals, conditional on θ, is given by:

$$(\theta_1 \ \theta_2 \ \theta_3 \ \theta_4)^t \sim \mathcal{N}(\boldsymbol{\theta}, \boldsymbol{\Sigma}); \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.8 & 0.7 & 0.5 \\ 0.8 & 1 & 0.3 & 0.6 \\ 0.7 & 0.3 & 1 & 0.4 \\ 0.5 & 0.6 & 0.4 & 1 \end{pmatrix}.$$

The "closest" algorithm induces $\mathcal{S}_i^{\mathcal{C}} = \{1, 2, 3\}$. Assume the "random" algorithm induces $\mathcal{S}_i^{\mathcal{R}} = \{1, 3, 4\}$. Posterior variances are $\mathrm{Var}[\theta|\{\theta_1, \theta_2, \theta_3\}] = 0.58$ for the "closest" algorithm and $\mathrm{Var}[\theta|\{\theta_1, \theta_3, \theta_4\}] = 0.68$ for the "random" algorithm. In this case, learning is better under the algorithm that the engagement-maximizer platform provides.

In general, learning is not monotone in neighborhood growth under the "closest" algorithm, for a fixed k. As seen in the example, it might be the case that adding a more similar neighbor increases learning due to the covariances between this neighbor and the others appearing in the feed. However, this phenomenon disappears as N_i becomes large, and learning becomes inexistent in the limit: the posterior variance for the user converges to σ^2, exactly the same as if she was alone. This is shown in Figure 1 and Figure 2, where we simulate learning and total utility for a user whose neighborhood grows from $N_i = k$ until $N_i = 5000$. We fix $\lambda = 0.5$, $\beta = 0.2$ and $k = 30$. The size of the feed is $k = 40$. We observe that, for the "closest" algorithm, learning non-monotonically decays as the neighborhood grows in size: the posterior variance grows and tends to σ^2. The expected posterior variance of the random algorithm is $\frac{\sigma^2}{k}$. There are some inicial periods for which learning under the "closest" algorithm might outperform learning under the "random" algorithm, but this tendence disappears relatively fast. Total utility increases concavely and we observe the importance of network effects: while the "random" algorithm induces a similar utility independently of the size of the neighborhood, the utility induced by the "closest" algorithm dramatically increases.

Proposition 4.2. *Under the "closest" algorithm, user i's learning becomes negligible as $N_i \to \infty$:*

$$\lim_{N_i \to \infty} Var[\theta|\{\theta_j\}_{j \in \mathcal{S}_i^{\mathcal{C}}}] = \sigma^2.$$

Corollary 4.3. *As N_i grows large, the probability that the "random" algorithm outper-*

forms the "closest" algorithm in terms of learning converges to 1:

$$\lim_{N_i \to \infty} \mathbb{P}\left[Var[\theta|\{\theta_j\}_{j\in\mathcal{S}_i^{\mathcal{C}}}] > Var[\theta|\{\theta_j\}_{j\in\mathcal{S}_i^{\mathcal{R}}}]\right] = 1.$$

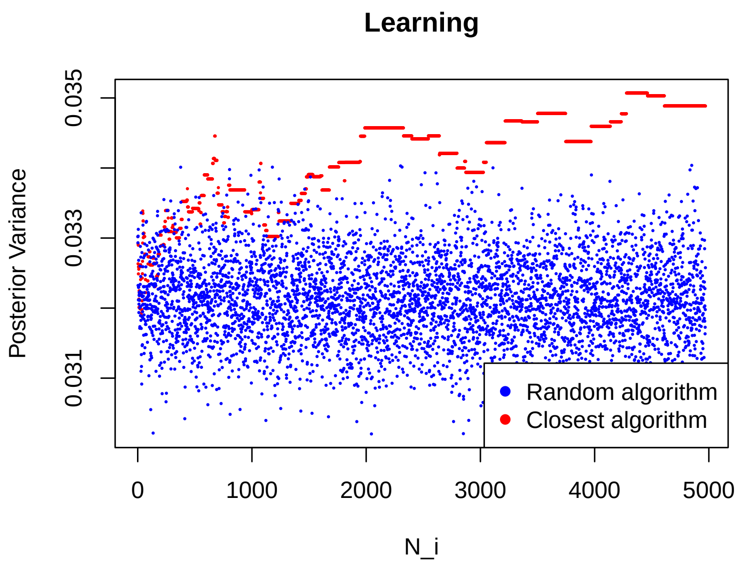

Figure 1: User's learning as the neighborhood size grows.

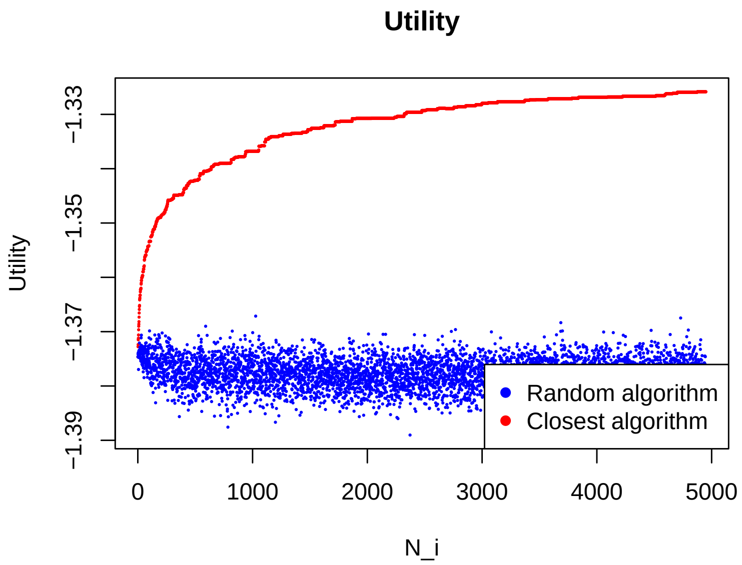

Figure 2: User's total utility as the neighborhood size grows.

Since the "closest" algorithm performs poorly in terms of learning and there is no learning at all in the limit, we focus on potential measures to improve the engagement-maximizing algorithm. A regulator could enforce the inclusion in the feed of neighbors selected using different criteria. We examine the effects on user utility of adding an extra user to a feed that has been already selected by the "closest" algorithm, when neighborhood size grows large. We consider three different possibilities: (i) a user that follows the "closest" algorithm, (ii) a user that is uncorrelated with any user in the feed, and (iii) a user that is maximaly and negatively correlated with every other user in the feed. We find that if conformity is not disproportionaly weighted compared to learning, the third measure is the most effective. It leads to maximal learning while the loss in terms of conformity is not too high. This measure could be referred to as "breaking echo-chambers", as it introduces a user in the feed who, in principle, holds opposite views to everyone else. While this measure might be implementable (requiring a regulator to enforce it), it may not be stable in the long-run. It is plausible that a user who is not fully rational would cut the link to a person who repeatedly expresses opposing views.

Proposition 4.4. *Given user i's feed of k neighbors selected using the "closest" algorithm as $N_i \to \infty$, the effects on total utility of adding an extra user are as follows:*

(i) *If the user is chosen according to the "closest" algorithm, both conformity and learning do not vary and are kept to 0 and -σ^2 respectively.*

(ii) *If the chosen user is uncorrelated with everyone else appearing in the feed, conformity becomes $\frac{-2\sigma^2}{k}$, and learning increases to $-\frac{\sigma^2}{2}$.*

(iii) *If the chosen user is maximaly and negatively correlated to everyone else appearing in the feed, conformity becomes $\frac{-4\sigma^2}{k}$ and learning increases to 0.*

Finally, it is better to add a negatively correlated user if and only if

$$\lambda < \frac{k}{4(1-\beta)+k}.$$

Even though there might not be a clear domination between the "closest" and the "random" algorithms for a social planner, they are both inferior to the "user optimal" algorithm regarding social welfare. We devote the next section to analyzing whether competition would make platforms implement the "user optimal" algorithm in equilibrium.

5 Platform Competition: Interoperability

Above, we have discussed the challenges of improving the "closest" algorithm, which limits the available tools for a regulator to enhance social welfare in a monopolistic market. Moreover, users would like the platform to implement the "user optimal" algorithm. This is no easy task, once we discard imposing it to the monopolist platform.

To circumvent the monopolistic issue, we propose utilizing platform competition. In theory, if users prefer the "user optimal" algorithm, any platform implementing it could attract the monopolist's user base if the monopolist does not change its algorithm. However, two intrinsic aspects of social media platforms complicate this situation and discourage competition: network effects and monopolist data advantage. Network effects, abstracting specific neighborhood considerations, indicate that the expected utility of being on a platform increases with the size of the neighborhood. Consequently, larger platforms find it easier to attract and retain users. Notably, the "random" algorithm lacks network effects, whereas personalized algorithms like the "user optimal" and "closest" algorithms exhibit them. Monopolist data advantage refers to the fact that when a user decides to leave the monopolist and join a competing platform, the new platform has no information about past interactions, making it unable to personalize the new user's feed initially. A consequence of these two properties is that users consider not only the algorithm's effects but also the benefits incumbents derive from their size and data. Incumbents may retain their user base while implementing socially inferior algorithms, deterring potential competitors.

Proposition 5.1. *For sufficiently large λ, both the "user optimal" algorithm and the "closest" algorithm feature network effects.*

The presence of network effects and monopolist data advantage hinders competition and leads to market concentration. Generally, the stronger these effects, the fewer platforms can remain profitable in the market. In social media platforms, this manifests as a "winner-takes-all" scenario, resulting in a monopoly. Even if a potential entrant chooses to adopt the "user optimal" algorithm, the incumbent would counteract by adopting the same algorithm. Due to network effects, users would not join the entrant, prompting the entrant to prefer staying out of the market. Major monopolists like Facebook or Twitter have heavily campaigned to demonstrate algorithmic changes or their efforts against fake news when public opinion started to doubt, aiming to deter potential competitors, as evidenced in Horwitz et al. (2021).

We propose interoperability as a solution to this problem. Interoperability refers to complete interaction between different platforms, eliminating network effects. Two platforms become interoperable when their users can interact with each other and data is shared. Hence, an entrant platform could use the whole population to provide each user with the personalized feed it desires: messages posted in the monopolist platform could be displayed in the entrant platform and viceversa. Moreover, data would be shared, and monopolist data-advantage also disappears. This forces both platforms to play "user optimal" algorithm in equilibrium, in the spirit of à la Bertrand competition.

Interoperability has been enforced in industries such as cell phones and email. Naturally, the level playing field created by interoperability disadvantages platforms with significant network effects, as consumer adoption decisions are no longer influenced by size. Conversely, smaller platforms would fear losing if they competed "for the market" and thus prefer interoperability to be able to compete "in the market" (Belleflamme and Peitz, 2019).

6 Conclusion

We have built a model of communication and learning through personalized feed. An engagement-maximizing monopolist platforms has no incentives to take social learning into account, so it chooses to show users the messages of those neighbors who are the most similar to them. This is a suboptimal outcome in social welfare terms. We show that competition is not enough to incentivize platforms to implement the "user optimal" algorithm in equilibrium. Nevertheless, if interoperability is enforced, competing platforms will select the socially optimal algorithm in equilibrium. Interoperability is prevalent in certain telecommunications markets, such as cell phone and email services. However, implementing it in social media platforms may be more challenging since user private information is much more substantial, and privacy concerns may arise.

Our model can be further developed to investigate certain aspects that this paper does not address. Specifically, future research can explore why and how monopolies arise by constructing a model of dynamic competition with heterogeneous platforms.

References

Abreu, Luis and Doh-Shin Jeon, "Homophily in social media and news polarization," working paper, available at *https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=3468416*, 2019.

Acemoglu, Daron and Asuman Ozdaglar, "Opinion dynamics and learning in social networks," *Dynamic Games and Applications*, 2011, *1* (1), 3–49.

_ , _ , and James Siderius, "Misinformation: Strategic sharing, homophily, and endogenous echo chambers," Technical Report, National Bureau of Economic Research 2021.

Agarwal, Saharsh, Uttara M Ananthakrishnan, and Catherine E Tucker, "Deplatforming and the Control of Misinformation: Evidence from Parler," working paper, available at *https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4232871*, 2022.

Allcott, Hunt and Matthew Gentzkow, "Social media and fake news in the 2016 election," *Journal of Economic Perspectives*, 2017, *31* (2), 211–36.

Belleflamme, Paul and Martin Peitz, "The competitive impacts of exclusivity and price transparency in markets with digital platforms," working paper, available at https://econpapers.repec.org/paper/corlouvco/2019019.htm, 2019, (2019019).

Benzell, Seth and Avinash Collis, "Regulating Digital Platform Monopolies: The Case of Facebook," working paper, available at *https://www.aeaweb.org/conference/2023/program/paper/SNNd74ni*, 2022.

Bernheim, B Douglas, "A theory of conformity," *Journal of Political Economy*, 1994, *102* (5), 841–877.

Biddle, Sam, "Facebook uses artificial intelligence to predict your future actions for advertisers, says confidential document," *The Intercept*, 2018, *13* (04), 2018.

Chamley, Christophe, *Rational herds: Economic models of social learning*, Cambridge University Press, 2004.

Cialdini, Robert B and Noah J Goldstein, "Social influence: Compliance and conformity," *Annual Review of Psychology*, 2004, *55* (1), 591–621.

DeGroot, Morris H, "Reaching a consensus," *Journal of the American Statistical Association*, 1974, *69* (345), 118–121.

DeMarzo, Peter M, Dimitri Vayanos, and Jeffrey Zwiebel, "Persuasion bias, social influence, and unidimensional opinions," *Quarterly Journal of Economics*, 2003, *118* (3), 909–968.

Ellman, Matthew and Fabrizio Germano, "What do the papers sell? A model of advertising and media bias," *Economic Journal*, 2009, *119* (537), 680–704.

Evens, Tom, Karen Donders, and Adelaida Afilipoaie, "Platform policies in the European Union: Competition and public interest in media markets," *Journal of Digital Media & Policy*, 2020, *11* (3), 283–300.

Franck, Jens-Uwe and Martin Peitz, "Market Power of Digital Platforms," *Available at SSRN, https://ssrn.com/abstract=4205570*, 2022.

Gottfried, Jeffrey and Elisa Shearer, "News use across social media platforms 2016," 2019.

Habib, Hussam, Maaz Bin Musa, Fareed Zaffar, and Rishab Nithyanand, "To act or react: Investigating proactive strategies for online community moderation," *arXiv preprint arXiv:1906.11932*, 2019.

Hartigan, John A, *Bayes theory*, Springer Science & Business Media, 1983.

Horwitz, Jeff et al., "The Facebook Files," *The Wall St Journal. Available online at: https://www.wsj.com/articles/the-facebook-files-11631713039(accessedNovember01,2021)*, 2021.

Hu, Lin, Anqi Li, and Xu Tan, "A rational inattention theory of echo chamber," working paper, available at https://arxiv.org/pdf/2104.10657.pdf, 2021.

Hwang, Elina H and Stephanie Lee, "A nudge to credible information as a countermeasure to misinformation: Evidence from twitter," working paper, available at https://ssrn.com/abstract=3928343, 2021.

Jackson, Matthew O, Suraj Malladi, and David McAdams, "Learning through the grapevine and the impact of the breadth and depth of social networks," *Proceedings of the National Academy of Sciences*, 2022, *119* (34).

Jadbabaie, Ali, Pooya Molavi, Alvaro Sandroni, and Alireza Tahbaz-Salehi, "Non-Bayesian social learning," *Games and Economic Behavior*, 2012, *76* (1), 210–225.

Kosinski, Michal, David Stillwell, and Thore Graepel, "Private traits and attributes are predictable from digital records of human behavior," *Proceedings of the National Academy of Sciences*, 2013, *110* (15), 5802–5805.

Kranton, Rachel and David McAdams, "Social networks and the market for news," working paper, available at https://sites.duke.edu/mcadams/files/2020/05/KrantonMcAdamsWP.pdf, 2020.

Lauer, David, "Facebook's ethical failures are not accidental; they are part of the business model," *AI and Ethics*, 2021, *1* (4), 395–403.

Levy, Ro'ee, "Social media, news consumption, and polarization: Evidence from a field experiment," *American Economic Review*, 2021, *111* (3), 831–870.

Loomba, Sahil, Alexandre de Figueiredo, Simon J Piatek, Kristen de Graaf, and Heidi J Larson, "Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA," *Nature Human Behaviour*, 2021, *5* (3), 337–348.

Molavi, Pooya, Alireza Tahbaz-Salehi, and Ali Jadbabaie, "A theory of non-Bayesian social learning," *Econometrica*, 2018, *86* (2), 445–490.

Mosleh, Mohsen, Cameron Martel, Dean Eckles, and David G Rand, "Shared partisanship dramatically increases social tie formation in a Twitter field experiment," *Proceedings of the National Academy of Sciences*, 2021, *118* (7), e2022761118.

Mostagir, Mohamed and James Siderius, "Learning in a post-truth world," *Management Science*, 2022, *68* (4), 2860–2868.

_ and _ , "When do misinformation policies (not) work?," *Available online at: https://siderius.lids.mit.edu/how-do-we-stop-misinformation/ (accessed January 26, 2023)*, 2023.

_ and _ , "When should platforms break echo chambers?," *Available online at: https://siderius.lids.mit.edu/wp-content/uploads/sites/36/2023/01/quarantine-v6.pdf (accessed January 26, 2023)*, 2023.

Mudambi, Maya and Siva Viswanathan, "Prominence Reduction versus Banning: An Empirical Investigation of Content Moderation Strategies in Online Platforms," 2022.

Mueller-Frank, Manuel and Claudia Neri, "A general analysis of boundedly rational learning in social networks," *Theoretical Economics*, 2021, *16* (1), 317–357.

_ , Mallesh M. Pai, Carlo Reggini, Alejandro Saporiti, and Luis Simantujak, "Strategic Management of Social Information," working paper, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4434685, 2022.

Pennycook, Gordon and David G Rand, "Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning," *Cognition*, 2019, *188*, 39–50.

_ and _ , "The psychology of fake news," *Trends in Cognitive Sciences*, 2021, *25* (5), 388–402.

Popiel, Pawel, "Addressing platform power: The politics of competition policy," *Journal of Digital Media & Policy*, 2020, *11* (3), 341–360.

Przybylski, Andrew K, Kou Murayama, Cody R DeHaan, and Valerie Gladwell, "Motivational, emotional, and behavioral correlates of fear of missing out," *Computers in Human Behavior*, 2013, *29* (4), 1841–1848.

Reuter, Jonathan and Eric Zitzewitz, "Do ads influence editors? Advertising and bias in the financial media," *Quarterly Journal of Economics*, 2006, *121* (1), 197–227.

Sagioglou, Christina and Tobias Greitemeyer, "Facebook's emotional consequences: Why Facebook causes a decrease in mood and why people still use it," *Computers in Human Behavior*, 2014, *35*, 359–363.

Silverman, Craig, "This analysis shows how viral fake election news stories outperformed real news on Facebook," *BuzzFeed,* available at https://newsmediauk.org/wp-content/uploads/2022/10/Buzzfeed-_This_Analysis_Shows_How_Viral_Fake_Election_News_Stories_Outperformed_Real_News_On_Facebook_-_BuzzFeed_News.pdf, November 16, 2016.

Solon, Olivia, "Facebook's failure: Did fake news and polarized politics get Trump elected?," *The Guardian,* available at *https://www.theguardian.com/technology/2016/nov/10/facebook-fake-news-election-conspiracy-theories*, November 10, 2016.

Sunstein, Cass R, "A prison of our own design: divided democracy in the age of social media," *Democratic Audit UK*, 2017.

Törnberg, Petter, "Echo chambers and viral misinformation: Modeling fake news as complex contagion," *PLoS one*, 2018, *13* (9).

A Example

Here we present the feeds that user i, with a neighborhood (Figure 3) composed by 13 users ($N_i = 13$), would receive under the "closest" algorithm (Figure 4), the "random" algorithm (Figure 5), and the "user optimal" algorithm (Figure 6). Note that we assume $k = 5$.

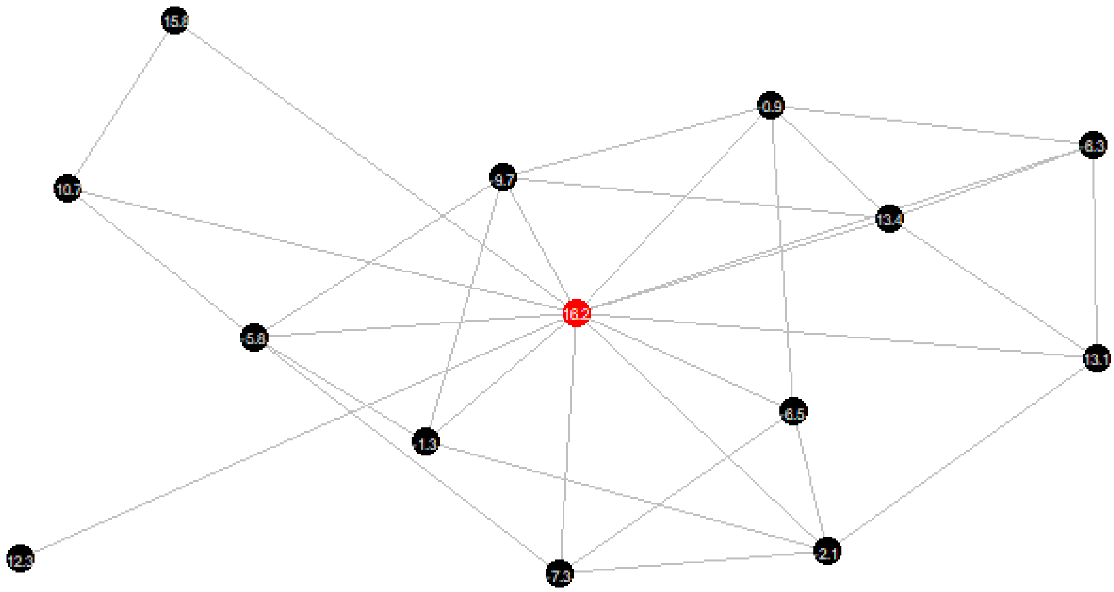

Figure 3: User i's neighborhood.

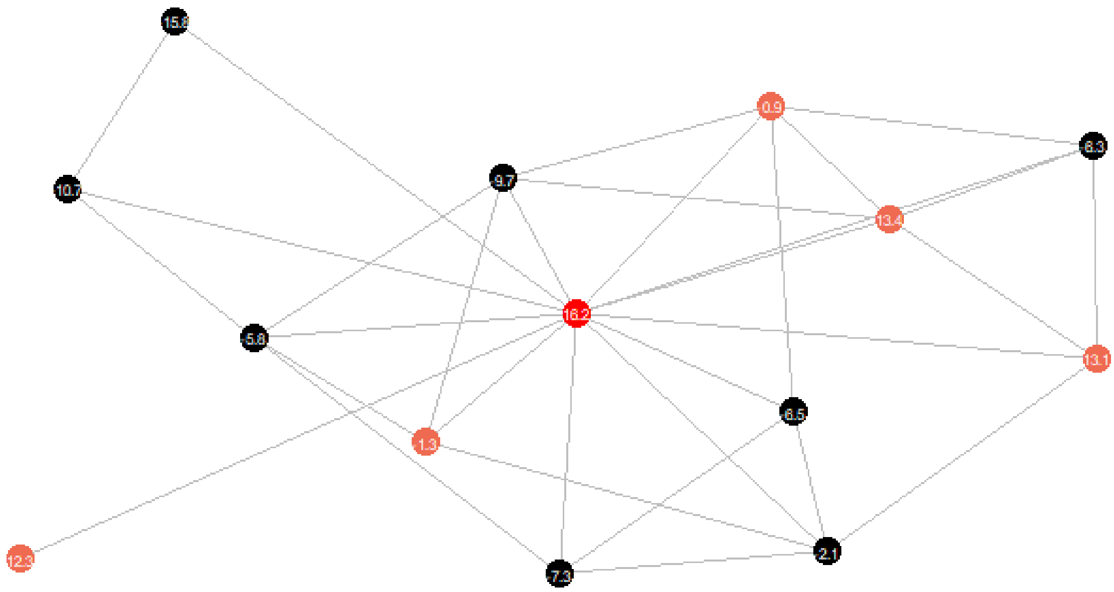

Figure 4: "Closest" algorithm feed.

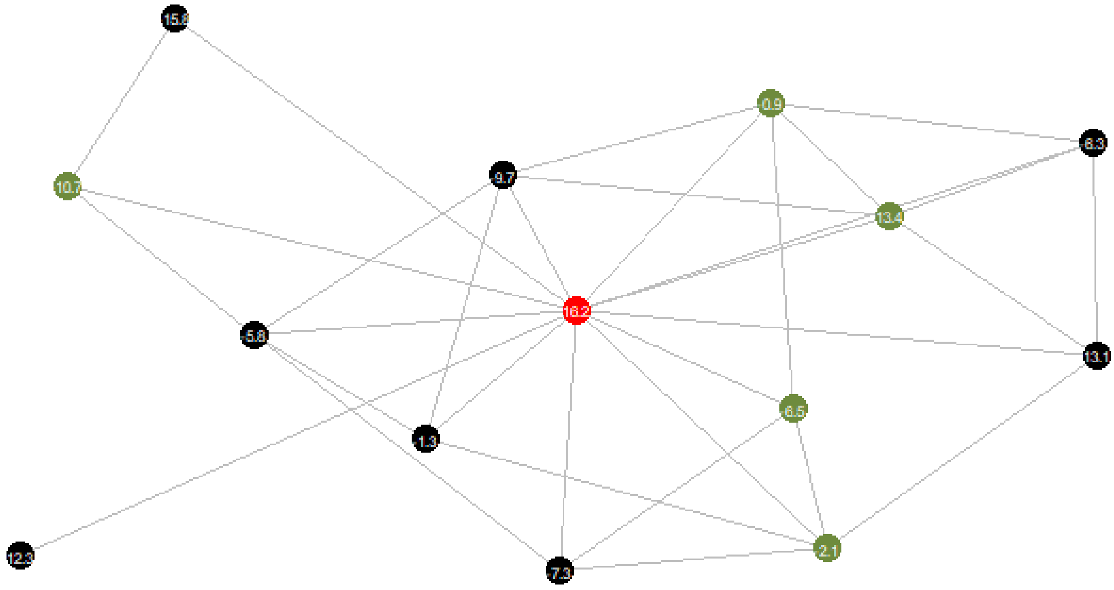

Figure 5: “Random” algorithm feed.

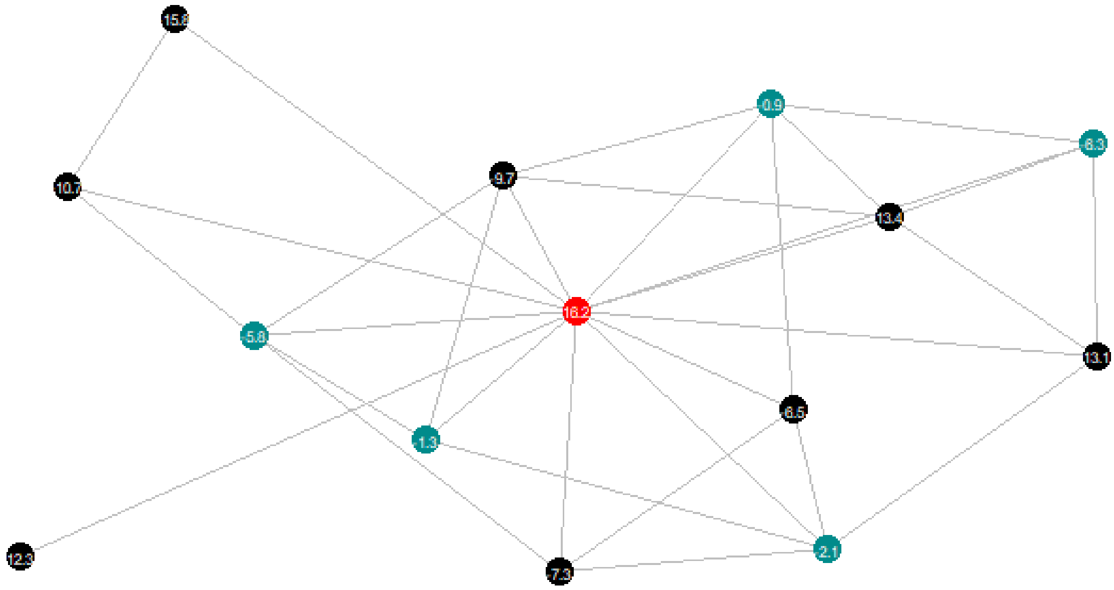

Figure 6: “User optimal” algorithm feed.