Queen's Economics Department Working Paper No. 1293

Getting the Right Spin: A Theory of Value of Social Networks

Pier-AndrÃľ Bouchard St-Amant
Queen‘s Economics Department

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

9-2012

Getting the Right Spin: A Theory of Value of Social Networks

Pier-André Bouchard St-Amant*

September 2013

Summary

I examine the problem of maximizing the spread of information in a context where users of a network decide which piece of information is shared. A company thus provides initial information to some users and they then choose what to share to their neighbours. These actions of sharing and choosing produce the characteristics of word-of-mouth advertising over time. I then answer the two following questions: what is the best word-of-mouth campaign that the company can perform and second, what is the value of such a campaign? The optimal solution can be understood as a Nash Equilibria that maximizes the concentration of the initial information to a small group of users. Such solution contrasts with standard measures of user influence and I show that they can sometime be seriously misleading. I provide an exact solution for a wide class of generic network topologies and an algorithm to compute it in polynomial time.

Keywords: Network Economics, Network Value, Word-Of-Mouth Advertising.

JEL Classification Numbers: L1, D83, D85.

*Queen's University, Department of Economics, 94 University Ave., Kingston, Ontario, Canada K7L 3N6, pabsta@econ.queensu.ca. The Social Sciences and Humanities Research Council (SSHRC) provided financial support for this research. I am thankful to Vincent Boucher, Frédéric Deroïan and James Bergin for their helpful comments. I am also thankful for three anonymous referees who provided meaningful comments and remarks. The most up-to-date version of this draft can be found at http://pabsta.qc.ca/en/research/vsn .

1 Introduction

The diffusion of information on a social network is an important problem in economics and the other social sciences. Studies on specific networks have been done to understand the spread of rumours about a stock, or "herding" [4] or if the heterogeneity in information sources leads to the convergence of ideas or actions [13, 11]. Some studies have looked at how personal attributes affect the speed of the diffusion of ideas [12].

This importance goes beyond being a simple transmission mechanism. Some studies of how voting is "contagious" on a network suggest that the network structure is more important than the message itself:

> The effect of social transmission on real-world voting was greater than the direct effect of the messages themselves, and nearly all the transmission occurred between "close friends" who were more likely to have a face-to-face relationship. These results suggest that strong ties are instrumental for spreading both online and real-world behavior in human social networks. [5]

Diffusion on a network can also induce cumulative or "ripple" effects that are detrimental. For instance, [9] study how the network structure influences the propagation of business shocks while [1] show how the network structure can amplify business shocks at the firm level. A proper understanding of diffusion patterns can thus help us reduce aggregate fluctuations.

This paper also studies diffusion in the context of **word-of-mouth advertising**. It takes the viewpoint that word-of-mouth is a deliberate choice made by users on the network. It is a joint realization of a monopolistic company, which seeks to maximize profits by advertising, and users, who choose what piece of information they share with other in their connections. The company thus provides initial information - a **seed** - to users and, depending on the information provided, users choose what they share with their neighbours. The information shared can then be chosen again by other users in the neighbourhood, and so on. Hence, information becomes viral because it is selected repeatedly over time. Consequently, an analysis of influence based solely on the network structure can be significantly misleading. Some users might appear influential based on their position in the network, but they might never be selected as an information source.

This paper contributes to the literature through three features that are not found together in previous articles. First, the model embeds this discussed notion of information selected by users. The observed diffusion patterns are thus the outcome of joint decisions made by both the company and the users on the network. Second, the model does not depend on a particular network formation process, or a network topology, to characterize the optimal solution. It can thus be applied to *any* observed[1] network structure. Third,

[1] I discuss in the applications an extension where the structure is observed with given probabilities. All the results found in this paper hold with natural modifications.

the task of finding the optimal solution is of low dimensionality. An exact solution can thus be found and applied on large scale networks. Hence, I provide an algorithm that can be computed in polynomial time. These features together make the model easily adaptable for an empirical application with the well-known discrete choice model.

1.1 Related Literature

Some papers study whether diffusion leads to the convergence of actions or ideas in the long run. In these models, agents process information from their neighbours according to a given rule. Individuals in [17] use Baye's rule while those in [12] simply average information received (as in [13]). The main insight of those papers shows under which conditions the agents' actions or ideas converge in the long run.

A recurrent approach of these theoretical papers is that the transmission of information bears no connection to decisions by users on the networks. It is either the result of a random connection between two users or simply a mechanical rule (Baye's rule, averaging) on the information available. This seems at odd with common experience on social networks like Twitter or Facebook. Users choose to share (or "retweet") the information they want according to their interest. The idea of averaging or processing information according to Baye's rule makes little sense in that context. What is diffused is the result of a choice made by users.

Other papers look at profit maximization, or how a firm can exploit how information is shared between agents on the network. Schiraldi and Liu [18] explore whether a firm should launch a product into different markets sequentially or simultaneously. They find that the firm can increase its profit by manipulating the order of the launch sequence order. Galeotti & Goyal [11] also look at profit-maximizing word-of-mouth and assume that the firm can only observe the probability that two agents share information. In Campbell [7], profit maximization is explored on a random graph where connections in the network depend on the price of the object for sale. As a general rule, these models find that the investment of initial information in nodes increases if profits increase with word-of-mouth advertising.

These papers rely, however, on a particular network topology to derive their results: either because it is assumed that the complete network structure is unknown to the firm or to make the mathematics tractable. In the context of econometric research, or for companies who know the network structure (like Facebook or Twitter), these approaches give few insights and needs to be complemented with a theoretical characterization on a fully known network structure.

The purest version of the word-of-mouth problem is known as the "set covering problem" (see [8]) and is shown to be cursed with dimensionality[2] (or NP-hard). The third

[2]To illustrate, consider the following problem. There are n users on a network, and they receive a binary signal, 0 or 1. The users reproduce the given signal only if they receive it from a neighbour. If a

strand of the literature thus focuses on algorithms to find an approximation of the optimal solution. When the full information on the network is used, papers usually focus on fast algorithms that are guaranteed to be within a given range of the optimal solution [15]. Some papers look for such algorithm on a particular probabilistic model of diffusion [6, 15, 10]. Others look at an algorithm for a near-optimal pricing strategy given a probabilistic process of diffusion [2]. While these algorithms have their practical use, they remain approximations that are generally guaranteed to be within 63% of the optimal solution.

My contribution to the literature is to provide a model that embeds decisions on information selection, is fully general in its application and yields an algorithm that can find an exact solution. The observed patterns of diffusion are thus the outcome of profit maximization and the choices of users. The solution includes the short-run as well as the long-run outcome dynamics. I finally show that a consequence of the model is that what is know to be the "greedy" algorithm that leads to an optimal solution.

The rest of this paper is organized as follows. Section 2 presents the model itself, its main implications and the solution. In section 3, I present an algorithm to compute the optimal solution. Section 4 presents some applications examples. I show, in particular, how standard measures of influence can be misleading when users choose what to share. I also show how the model can be ported to an econometric framework and how to adapt the framework to partially known networks. I also show how time influences the optimal solution. A brief conclusion follows.

2 The Model

I use a directed graph $G(V, E)$ to model the relationship between users. V is a set of of vertexes of cardinality $|V|$. I refer to vertexes as **users** or **agents**. Likewise, $E \subseteq V \times V$ is the set of directed $|E|$ edges between users. I use the convention that an edge $e_{ij} \in E$ means that user j is a is **source of information** for user i. The set of sources of information for user i, or his **choice set of sources** is his inbound neighbourhood $\eta_i \equiv \{j \in V : e_{ij} \in E\}$.

Each of these sets is assumed to be fixed. This constrains users to choose within the set of friends each user already has. It is as if the advertising campaign is for the short-run, where the social capital — the structure of the network — is fixed. Hence, users can draw information only from sources they have already developed.

---

company has to choose the initial vectors of zeros and ones at a linear cost that maximizes the discounted sum of ones through time, there are 2^n possible candidate solutions. Although n is finite, on a network of the size of Facebook, there are $2^{300\times10^6}$ solutions to check. If it takes 10^{-9} seconds to compute one solution, the calculation would require a processing time of roughly $10^{90\times10^6}$ seconds. In comparison, the age of the universe is estimated to an order of 10^{17} seconds. On large-scale networks, a fast algorithm is thus of the utmost importance.

In this paper, I will use Figure 1 as a simple example of a possible network. In the Figure 1(a), there are five users (labelled by the index i ranging from one to five). Each directed edge represents a possible choice of source of information. For instance, user $i = 2$ has two sources of information, namely users $j = 1$ and $j = 5$ and so $\eta_1 = \{1, 5\}$.

[Figures 1(a) and 1(b) about here]

What drives the selection of information by users is a real number that I call a **signal**. I denote these signals s_{it}, which denotes a signal shared by user i at time t. Likewise, $s_t = [s_{1t}, s_{2t}, \ldots, s_{|V|t}]'$ denotes the whole vector of signals at time t. A signal is a simple device that measures users' valuation of the information being shared. Throughout the paper, I assume there are discussions about only one product and so there is a single signal. This is done for simplicity, as the case with multiple products is a trivial extension.

The exact meaning of "valuation of signals" depends on how the underlying information about the product is modelled. Three examples follow:

1. A product can be modelled as a vector in $\mathbb{R}^n$, where each dimension reveals something about the product (quality, color, probability of breaking, etc.). One piece of information on the product then represents a sigma-field with a measure on such vector, describing the likelihood of a particular realization of the object. If agents are endowed with strict preferences u over these sigma-fields, the signal represents the level of utility given by such a sigma-field. The choice is thus governed by how useful the information is to the user.

2. The signals are simply the econometric equivalent of the previous example, that is $s_{it} = X\beta + \epsilon_{it}$, where X is a vector of observed characteristics on the product, β is the (linear) weight given to each characteristic and ϵ_{it} is a "utility shock" following a random process.

3. The signal can also represent the evolution of demand, or the informativeness of a signal. Imagine that the firm has a good-quality product, but such quality is unknown to the users. Their priors are that there is a 50% chance that the product is of good quality and a 50% chance that it is not. The price of the product is set higher than 0.5 and the value of the product is one if it is of good quality and zero otherwise. Users can learn the quality about the product if they buy it, or if they see one of their neighbours buy it. Users will want to buy a product only if it is of good quality, thus they will do so only if they see at least one of their neighbours buy it. Hence, the signal chosen can be represented as the highest signal in the neighbourhood.

4. Signals are simply a measure of the "quantity of information" the company provides.

Users on the network have the same valuation of information. However, they do not value their neighbours in the same fashion. For each of its sources, a user has a **perception**, or a **bias** $b_{ij} \in \mathbb{R}$ towards it. In other words, users put more valuation on some given neighbours than on others. This is so either because they trust them, or because they share similar affinities. If all perceptions are equal to zero, then all sources are seen as neutral and equivalent to each other. These biases simply encode that we generally value some sources of information more than others.

For the purpose of this paper, these perceptions are fixed in time. This is done to derive a tractable analytical solution and this means that the relationship between users does not change over the course of the word-of-mouth campaign. In Figure 1(b), perceptions are depicted along the edges. For instance, user $i = 1$ has a perception of 3 towards user $j = 2$ ($b_{12} = 3$) while it has a perception of 4 regarding user $j = 3$ ($b_{13} = 4$). In this example, all perceptions are positive but nothing prevents some of them from being negative. As we shall see, what matters is the relative difference between each of the sources. As in consumption theory, decisions are based on comparing alternatives. Hence, perceptions are useful as an ordinal concept.

2.1 The Reaction Function and the Law of Motion of then Network

Signals represent how valuable the information is, so users choose to carry what they find the most valuable. In other words, the signal shared, or the **reaction function** to signals is defined by the following:

$$s_{it+1} \equiv \max_{j \in \eta_i} \left(s_{jt} + b_{ij} \right).$$

This reaction function means that a user carry the most valuable signal given the perceived valuation they have about their sources. In the examples of what signals might encode, the reaction function has the following interpretation: in the first two cases, it carries over the signals with the highest level of utility; in the third case, an informative signal is carried over directly if a user can observe from his or her neighbours the quality of the product. It represents the evolution of beliefs (and thus sales) about the product over the network. In the last case, the signal carried over is the most informative one.

When applied to the whole vector of signals, I denote $s_{t+1} = G(s_t)$ the **law of motion** on the whole network. I also use $s_{t+n} = G^n(s_t)$ to denote that the law of motion is applied n times.

Using again the example in Figure 1(b), users respond to information according to the following:

$$\begin{aligned} s_{1t+1} &= \max(s_{4t}, s_{5t} - 100), && (1)\\ s_{2t+1} &= \max(s_{1t}, s_{5t} - 100), && (2)\\ s_{3t+1} &= \max(s_{2t}, s_{5t} - 100), && (3)\\ s_{4t+1} &= \max(s_{3t}, s_{5t} - 100), && (4)\\ s_{5t+1} &= \max(s_{2t} + 1). && (5) \end{aligned}$$

If in time t, the vector of signals s_t is given by $[0, 0, 0, 0, 0]$, the law of motion will yield the vector $[0, 0, 0, 0, 1]$ in time $t + 1$. When all users have made their choice, the vector of signals s_{t+1} becomes the next set of signals and users come back in $t + 1$ to perform another selection.

Because users have fixed perceptions and these perceptions are additive, a counterintuitive consequence of the model is that signals can continuously increase (or decrease). For instance, this would happen if all perceptions are positive on each edge of the network. This consequence is however a technical simplicity of no importance. What matters is the outcome based on such valuation, not the valuation in itself. Like utility valuation, what is important is the relative difference between the possible choices. As we shall see, the company will seed signals proportionally to the number of times a signal is selected, which is the valuable outcome of the network of users. All in all, the company will derive a finite value of this reaction function.

This modelling choice implies that all users agree that information about the product is valuable. However, they disagree on the degree, through their perceptions. And since the information is seeded by the the company, this means that the information is also valuable from a sales perspective. So this paper assumes that both users and the firm think it is a good idea to spread information around.

With standard mathematics, this law of motion is nonlinear. However, I make the case that this law of motion is linear in a different algebra. I use this algebra to find properties of the model that are easily expressed in terms of linear first-difference equations. I describe this in details when solving the model.

2.2 The Company's Objective

The company is a monopolist that maximizes profits. It does so by selling viral exposure to a third party at price p per unit. The demand for such exposure is rooted in the idea that advertising increases sales (see [19] for a review).

Such a company wants the highest possible levels of valuation to reach each user. For a unit of valuation reaching a user at time t, the third party is willing to pay a unit price of $p\beta^t$. This means that it values current exposition more than future exposition as it discounts time at the rate $\beta \in [0, 1)$. Now, for an initial seed $s_0 \in \mathbb{R}^{|V|}$ and the resulting sequence $\{s_t\}_{t=0}^{\infty}$ induced by the law of motion, the **income of the company** is given by the following:

$$I(s_0, p) \equiv p \sum_{t=0}^{\infty} \beta^t \sum_{i \in V} s_{it},$$
$$\text{s.t. } s_t = G^t(s_0) \; \forall t > 0.$$

This income function assumes that the campaign runs forever as t is allowed to go to infinity. As we shall see, performing the analysis in finite time would not change the insights as long as the number of periods is larger than a threshold t^* that is smaller than the largest directed diameter of the graph.

The company has only one instrument at its disposal, that is the initial vector of signals, s_0, reflecting the first information each consumer gets about the product. Afterward, the evolution of signals is governed by the choices made by users in the network.

To produce a given signal signal s_{i0}, the firm must bear a quadratic cost $\frac{s_{i0}^2}{2}$. Hence, the total cost of the seed is $\sum_{i \in V} \frac{s_{i0}^2}{2}$. This assumption of quadratic costs reflects the idea that it becomes harder to produce signals with a higher valuation to users. This discussion leads to the following problem for the company:

Definition 1 (Company's problem)**.** *Let $\boldsymbol{0} \in \mathbb{R}^{|V|}$ denote the zero vector. The company then seek a seed of signals $s_0 \in \mathbb{R}|V|$ satisfying:*

$$\begin{aligned} \arg\max_{s_0} \Pi(s_0, p) &\equiv I(s_0, p) - I(\boldsymbol{0}, p) - c(s_0) \\ &s.t. \\ &s_{t+1} = G(s_t) \; \forall t \\ &c(s_0) = \sum_{i \in V} \frac{s_{i0}^2}{2} \end{aligned}$$

Denote s_0^* the solution[3] to this problem.

Profits are calculated as the difference between the net income generated by the chosen signal and the costs to build that signal. The net income is the difference between the value of the chosen campaign and the value of a campaign where the company does not seed any signal. As this value will be generated even if the company performs no action, it cannot charge for it. The component $I(\mathbf{0}, p)$ can thus be thought of as the social value of the network with respect to the product.

[3] One can easily show that the function is well defined, continuous and that s_0 is necessarily chosen in a compact set. It thus has a maximum.

2.3 The Word-of-Mouth Advertising Game

The model can be described as a repeated game with $|V| + 1$ players. The additional player is the firm with actions and payoff described in definition 1. That player plays only once at $t = 0$. The $|V|$ other players are users on the network, each with a fixed strategy space η_i. For a given action a_{it} which is the choice of a source in η_i, the pay-off at is given by $s_{jt} + b_{ij}$. The optimal action at each period is then given by the reaction function $\max_{j\in\eta_i} (s_{jt} + b_{ij})$. A **word-of-mouth campaign** is then the result of a **strategy profile** $(s_0, \{a_{it}\}_{t=0}^{\infty} \ \forall i)$ between the firm and the users on the network.

These choices a_{it} made by users in time t implicitly define the payoffs in $t + 1$. For instance, if user i plays action $a_{it} = j \in \eta_i$, the payoff associated with choosing source i in $t + 1$ is given by $s_{it+1} = s_{jt} + b_{ij}$. Going one step further, if s_{it+1} is chosen by another user, say user k, it will share a valuation $s_{kt+2} = s_{it+1} + b_{ki}$. By the previous token, this implies that $s_{kt+2} = s_{jt} + b_{ki} + b_{ij}$. Applying this recursive argument up to $t = 0$ shows that any signal s_{it}, at any given time, is the sum of a component of the original seed s_{k0} and a the sum of all perceptions along the path of selections. Such path of selection made this original seed travel in the time-node space from user k at the initial period to user i in period t. This result is proved formally in lemma 1 in Appendix A.1.

The solution concept is best expressed in terms of the number of times each original signal is chosen indirectly at a given time. I denote $k_{it}(a_{1t}, a_{2t}, \ldots, a_{|V|,t}) \in \{0, 1, \ldots, |V|\}$ as the number of users who end up choosing the original signal s_{i0} at time t through the recursion process expressed above. So k_{it} counts the **number of choices**, or how influential signal s_{i0} is at time t. These numbers define the collective outcome of the actions of each agent on the network and are sufficient numbers to express the optimal solution.

Sticking to the example in Figure 1(b), assume that the seed is given by: $s_0 = [s_{10}, s_{20}, s_{30}, s_{40}, s_{50}]' = [1, 2, 3, 4, 5]$. This means that $s_1 = [s_{40}, s_{10}, s_{20}, s_{30}, s_{20} + 2]'$. In period two, user three chooses user one as its source. This means he ends up choosing s_{20}, and since he is the only one choosing it, $k_{21} = 1$. However, signal s_{10} is now available to user three and user five since user two has carried it in the previous period. It is also chosen by these two, so $k_{10} = 2$.

2.4 Solving the Model

The solution concept is a Nash equilibrium that maximizes the concentration of the seed for a few user. I decompose such characterization in two parts. First, any optimal solution s^*_{j0} can be written as a discounted sum of selections over time:

$$s^*_{j0} = p \sum_{t=0}^{\infty} k^*_{jt} \beta^t. \tag{6}$$

A seed s_0^* generates a best response action profile $\{a_{it}^*(s_0^*)\}_{t=0}^{\infty}$ $\forall i$. That best response generates the sequence of $\{k_{it}^*\}_{t=0}^{\infty}$ $\forall i$. In turn, each of these measures of influence at each period generates an optimal seed s_0^*, as stated by 6, which maximizes profits.

What generates value for the firm is exposure. Hence, the number of times an original signal is carried over in the next period, as counted by k_{jt}, is what matters to measure exposure in the next period. Hence the discounted sum of the number of selections, multiplied by the price, represents the marginal value of a signal s_{j0}. Setting it equal to the marginal cost yields an optimal solution.

[Figures 2(a) and 2(b) about here]

The second step deals with the many Nash equilibria that satisfy such solutions. I show that the profit-maximizing one maximizes concentration of information over users. Such maximal concentration is in turn a function of the discounting factor (β). This second step requires the understanding of a network structure that emerges in the dynamics of signal selection, namely a *constrained star network*. Loosely speaking, a star network has all its nodes attached to a central group of nodes (see Figure 2(a)), the center of the star. In a constrained star, some nodes are not directly connected to the center, but are, however, connected to the center through a simple path of nodes (see Figure 2(b)). The notion of constrained star is defined formally as follows:

Definition 2 (Simple Path, Constrained Star Network)**.** *A simple path from i to j is a repetitionless sequence of directed edges $e_{ik_1}, e_{k_1k_2}, \ldots, e_{k_nj}$ that connects i to j.*

A constrained star network is a graph $G(V, E)$ with the following properties:

1. *A hub H: a subset of V such that for any two nodes $i, j \in H$, there exists at least one directed simple path from i to j.*
2. *For any other node $i \in V \setminus H$, there exists one directed path starting at some node $j \in H$ and going to i.*

The second step can then be described as follows. If a network $G(V, E)$ has at least one constrained star structure as a subset of its nodes and edges, then the average value of the seed at each node is equal to the discounted value of the selling price $\frac{p}{1-\beta}$. So what matters in finding the optimal solution is the spread of the value of initial signals around this average. Such a spread is maximized when all the seed is concentrated amongst a small group of users. When costs are quadratic, the concentration of the seed is equivalent to maximizing the variance amongst all Nash qquilibria.

The next proposition shows the idea that any possible solution to the company's problem must be a Nash equilibrium.

Proposition 1 (Candidate Solutions)**.** *Let $\tilde{S} = \{\tilde{s}_0, \{\tilde{a}_{it}\}_{t=0}^{\infty}\ \forall i\}$ be the set of candidate solutions. Then any $\tilde{s} \in \tilde{S}$ must satisfy:*

1. $\tilde{s}_{t+1} = G(\tilde{s}_t)$;

2. $k_{j,0} = 1\ \forall j$;

3. *Given* $\tilde{s}_{j0}$, *users make choices such that* $\tilde{k}_{jt}$ *holds* $\forall j, t$;

4. *Given these* $\tilde{k}_{jt}$, $\tilde{s}_0$ *solves:*

$$\tilde{s}_{j0} = p \sum_{t=0}^{\infty} \beta^t \tilde{k}_{jt}\ \forall j;$$

5. *If the network admits at least one constrained star topology, the average value of the initial signals is then given by:*

$$\frac{1}{|V|} \sum_{j \in V} \tilde{s}_{j0} = \frac{p}{1-\beta}.$$

Proof. See Appendix A.2 □

The Nash equilibrium is summarized in points 3 and 4 of the proposition. Given that $\tilde{s}_{j0}$ is the strategy of the company for user j, it must be that users on the network chooses that signal k_{jt} times at time t. Now given that users have chosen the signal j $\tilde{k}_{jt}$ times at time t, $\tilde{s}_{j0}$ must account for the marginal contribution $p\beta^t\tilde{k}_{jt}$ at time t. Summing over all periods yields the marginal income and thus, a candidate solution.

The fourth item of this proposition clarifies what matters in seeding a node. It is a measure of the discounted social influence of each node. Any solution has to account for the discounted number of users who will select the information. The more a source is selected the more it is influential. This contrasts with an analysis of influence based simply on the network structure.

It may seem counterintuitive that the optimal solution does not account explicitly for the value of perceptions. Perceptions are important given that they influence choices, so they are implicitly accounted for in each k_{jt}. To understand why they are not in the formula, imagine a small marginal change to $\tilde{s}_{j0}$ on profits. If such a change is small enough, it will not modify the selections (the order of valuations) of users on the network. Hence, the value of perceptions has no explicit role in the first order condition.

The last statement of proposition 1 says that the average value of all signals must equal the discounted value of the marginal income. This means that all candidate solutions on a network with at least one constrained star in it have the same average value in $t = 0$. What discriminates the elements in $\tilde{S}$ is the degree of variation amongst signals given to users.

In the example of Figure 1(b), more than one solution satisfies the conditions of proposition 1. Notice that the signal s_{50} is chosen at time $t = 1$ by other users only if $s_{50} - 100 > s_{j0}$. Thus, there are two cases. If the inequality holds, signal $s_{5,0}$ is chosen by all other users and such signal is then shared all subsequent periods. The solution can then be written as:

$$\begin{aligned}
\tilde{s}_{1,0} &= p, \\
\tilde{s}_{2,0} &= p + p\beta, \\
\tilde{s}_{3,0} &= p, \\
\tilde{s}_{4,0} &= p, \\
\tilde{s}_{5,0} &= p + 4p\beta + 5p\frac{\beta^2}{1-\beta}.
\end{aligned} \tag{High Concentration Eq.}$$

If the inequality does not hold, all signals are chosen from the closest neighbour on the periphery, and four different signals are shared every time. On a period of four iterations, user five receives each seed from user two. This means that the number of selection of each signal will change from one to two when it reaches user two. Thus, the solution is as follows:

$$\begin{aligned}
\tilde{s}_{1,0} &= p + p\frac{\beta}{1-\beta^4}\left[1 + 2\beta + \beta^2 + \beta^3\right] \\
\tilde{s}_{2,0} &= p + p\frac{\beta}{1-\beta^4}\left[2 + \beta + \beta^2 + \beta^3\right] \\
\tilde{s}_{3,0} &= p + p\frac{\beta}{1-\beta^4}\left[1 + \beta + \beta^2 + 2\beta^3\right] \\
\tilde{s}_{4,0} &= p + p\frac{\beta}{1-\beta^4}\left[1 + \beta + 2\beta^2 + \beta^3\right] \\
\tilde{s}_{5,0} &= p
\end{aligned} \tag{Low Concentration Eq.}$$

I refer to these solutions as the "high" and "low" concentration equilibrium. The high concentration solution holds when $s_{50} - 100 > s_{20}$ while the other one holds when it is not the case. Notice that $\tilde{s}_{5,0} - 100 > \tilde{s}_{2,0}$ holds only if $p(3\beta + 5\beta^2(1-\beta)^{-1}) > 100$, so this is a candidate equilibrium that belongs to $\tilde{S}$ if this is true. Otherwise, the low concentration equilibrium is the unique equilibrium. In both case, these candidate solutions are such that the seed sums to $\frac{p|V|}{1-\beta}$ since the network admits at least one constrained star.

If $p(3\beta + 5\beta^2(1-\beta)^{-1}) > 100$, which of the two equilibria should be picked? Intuitively, we see that the one with the highest concentration of the seed yields the highest profits because more information is shared through the process of contagion. But in this equilibria, there is a loss of value as information travels on the edges of negative perceptions. So, there might be a trade-off between the gain through the replication process and the loss of perceptions. I will show later that this is not the case.

2.5 Profits and Dynamics of the Model in the Long Run.

I now turn to the characterization of the optimal profits and the dynamics of the law of motion $s_{t+1} = G(s_t)$. The law of motion and profits/income are related as one must first characterize $I(0, p)$, the social value of the network, before calculating income and profits. I first introduce the main results and then show in details how to get them.

If the network admits at least one constrained star network, I show that it enters a periodic regime of period T after a finite transient time t^*. This means that after t^*, the law of motion G can be described by the following:

$$s_{i,t+T} = s_{i,t} + \lambda T \; \forall t > t^*.$$

The element λ is defined as the highest average of perceptions over all possible loops of the network (I describe this more below). The value T is pinned down by the greatest common divisor of all cycles with average value λ. Such a result implies. in turn, that users select the same neighbours after t^*, as they remain the most valuable sources.

This pattern of source selection generates a constrained star network. After t^*, the loop with the highest average fuels all the other users who connect to them. The pattern of connections for those users, the "branches," are constrained by the sources available. Hence, after t^*, highly perceived agents are selected over and over directly or indirectly.

These results give, in turn, a characterization of the profits formulation and the social value of the network. Such values can be interpreted as the value the advertiser gets for free, without any campaign design. It is the simple action of users talking about the product.

To prove these results, a bit of investment in mathematics is required. The law of motion is based on $s_{i,t+1} = \max_{j \in \eta_i}(s_{j,t} + b_{ij})$, which is nonlinear in standard algebra. However, it is linear in what is called a *tropical algebra* or a max-plus semiring. If we define the addition operator over numbers a and b as the maximum operator ($a \oplus b \equiv \max(a, b)$) and the multiplication operator over numbers a and b as standard addition ($a \otimes b \equiv a + b$), we then get the basic operations of a tropical algebra. If we define a transition matrix B by the elements $[B]_{ij} = b_{ij}$, the system can be written as $s_{t+1} = B \cdot s_t$, a first difference linear equation in such algebra. Thus, one can recover the essential tools of first difference equations like eigenvector/eigenvalue decomposition.

I introduce with more detail the main structure of linear systems on a tropical algebra in Appendix B. Readers looking for a complete and formal treatment should read the book by Bacelli and al. [3]. In the body of this article, I present only the important definitions and theorems to get the results. To begin with, I start by defining "highest average loops", which is the notion of **critical cycle**. Critical cycles pins down the increase of λ and the length of the periodic regime.

Definition 3 (Critical Cycles)**.** *Let C be the set of all cycles on $G(V, E)$. For any cycle*

$c \in C$ of length $|c|$, let

$$\rho(c) \equiv \sum_{e_{ij} \in c} \frac{b_{ij}}{|c|}$$

be the average value of perceptions and $C(G)$ be the set of all cycles in a network $G(V, E)$. A cycle c_c is critical if

$$c_c \in \arg \max_{c \in C(G)} \rho(c).$$

Denote $\lambda \equiv \max_{c \in C(G)} \rho(c)$.

This definition assigns a number to all cycles in the network, namely the average increase of the signal through perceptions over each cycle. The critical cycles are the loops with the highest average. Notice that by construction, any cycle is an hub.

In Figure 1(b), there are two cycles: $\{e_{25}, e_{52}\}$ and $\{e_{21}, e_{32}, e_{43}, e_{14}\}$. They have an average value of $\frac{-100+1}{2} = -49.5$ and $\frac{0+0+0+0}{4} = 0$. Hence, the periphery cycle has the highest average and is thus a critical cycle.

Critical cycles are important because they contain users who create the highest consistent increase of the signal over time.[4] They influence the valuation in the long run. Since users choose signals they most value, they will eventually be influenced by the users in critical cycles, those that have the highest perceptions.

Now, I state a result on tropical algebras to show the existence of the periodic regime. I start with the definition of eigenvectors and eigenvalues in the context of this model and then state the main mathematical result.

Definition 4 (Iterated Law of Motion, Eigenvalues and Eigenvectors)**.** *Let $s_{t+1} = G(s_t)$ be the law of motion associated with a network $G(V, E)$. Furthermore, let $B \cdot s_t$ be its representation in a tropical algebra. An eigenvalue-eigenvector pair (λ, v) is a solution to the equation:*

$$B^T \cdot v = \lambda^T \otimes v,$$

or, in standard algebra:

$$G^T(v) = v + T\lambda.$$

With these definitions in hand, one can show the following:

Theorem 1. *For any network $G(V, E)$ that admits at least one constrained star, then:*

1. *There exists a transient time $t^* < \infty$ after which the network enters a periodic regime of frequency T. T is given by the least common multiple of the length of all critical cycles while t^* is no longer than the longest directed path between any two points;*

[4] In the example, this increase is equal to zero.

2. *There exists $v_1, v_2, \ldots v_T$ eigenvectors such that $G^T(v_i) = v_i + T\lambda$. Each eigenvector corresponds to one period;*

3. *These vectors are formed by the columns of the matrix:*

$$[B^*]_{ij} \equiv \max([B^T - T\lambda]_{ij}, [B^{2T} - 2T\lambda]_{ij}, [B^{3T} - 3T\lambda]_{ij}, \ldots).$$

4. *In particular, define $[v^*]_i \equiv \max(v_{1i}, v_{2i}, \ldots v_{Ti})$, then v^* is the unique eigenvector on $G(V, E)$:*

$$G(v^*) = v^* + \lambda;$$

Proof. See [3]. In particular, section 3.2.4 pp 111-116, for the notion of eigenvectors and section 3.7 pp. 143-151 for the cyclicality and finiteness of transient time. □

The most stringent requirement is that there exists at least one cycle in the network. This means that there is a group of users who listen to each other. If there is no such cycle, then any word-of-mouth campaign will end in finite time t^*. In this case, all $k_{jt} = 0$ after t^* and the conditions for proposition 1 remains.

But if there is at least one cycle, the system will enter a periodic regime. The nodes that have no path connecting them to the cycle will have no long-run value, as if there is no cycle. Hence, they change very little in the analysis. Given the clustering features of most real networks, having at least one cycle in a network is a weak condition.

In the example of Figure 1(b), the only critical cycle has a value of 0, so $\lambda = 0$. There are four nodes on this cycle, so $T = 4$. The eigenvectors are given by $v_1 = [0, 25, 0, 0, 1]'$, $v_2 = [0, 0, 25, 0, 26]'$, $v_3 = [0, 0, 0, 25, 1]'$ and $v_4 = [25, 0, 0, 0, 1]'$. This means that $v^* = [25, 25, 25, 26]'$.

The eigenvalue measures the average increase of valuation due to perceptions of influential users (those on the critical cycle). Eigenvectors measure the increases, net of λ, due to choices made by users. The interpretation for the components of v^* is similar: it gives the maximal increase that will occur over all periods given the selections made by users.

The last proposition allows profits and income to be found for the optimal solution.

Proposition 2. *Let $G(V, E)$ admit at least one constrained star topology and consider $\tilde{s} \in \tilde{S}$. Then:*

1. *Any $\tilde{s}$ can be written as:*

$$\tilde{s}_{j0} = p \sum_{t=0}^{t^*} \beta^t \tilde{k}_{jt} + p \frac{\beta^{t^*+1}}{1 - \beta^T} \sum_{u=0}^{T-1} \beta^u \tilde{k}_{j,u},$$

where $\tilde{k}_{j,u}$ is the number of times signal $\tilde{s}_{j0}$ is chosen in the u-th period of the periodic regime;

2. The income generated by the network is given by:

$$I(\tilde{s}_0, p) = \sum_{i=0} (\tilde{s}_{i,0})^2 + p\lambda \frac{|V|\beta^2}{(1-\beta)^2} + p\frac{\beta^{t^*+1}}{1-\beta^T} \sum_{u=0}^{T-1} \beta^u \sum_{i \in V} (\tilde{s}_{iu} - \tilde{s}_{iu-1} - \lambda)$$
$$\cdots + p \sum_{t=1}^{t^*} \beta^t \sum_{i \in V} (\tilde{s}_{it} - \tilde{s}_{it-1} - \lambda)$$

3. Denote **0** *the zero vector. Then, profits are given by:*

$$\Pi(\tilde{s}_0, p) = \frac{1}{2} \sum_{i=0} (\tilde{s}_{i,0})^2 + p\frac{\beta^{t^*+1}}{1-\beta^T} \sum_{u=0}^{T-1} \beta^u \sum_{i \in V} (\tilde{s}_{iu} - \tilde{s}_{iu-1}) \dots$$
$$\cdots + p \sum_{t=1}^{t^*} \beta^t \sum_{i \in V} \left(\tilde{s}_{iu} - \tilde{s}_{iu-1} - \left[G^t(\boldsymbol{0}) - G^{t-1}(\mathbf{0})\right]_i\right).$$

Proof. See Appendix A.3. □

The first statement is a simple corollary of the periodic nature of the law of motion. Since after t^* the choices become periodic, the optimal solution can be written as a discounted sum of every period.

The second statement characterizes the total value of the network. It states that the optimal income is the sum of four components. The first term is the income that stems from the value of $\tilde{s}^*_{j0}$. The second term is the value generated by users on critical cycles. Because users look for higher signals, all of them eventually connect to these signals through a path of selections. The third term is the value of perceptions that is implied by the chosen campaign. It measures the increase of value from perceptions in the periodic regime. The last term is the value of perception generated by transitionary dynamics (from $t = 0$ to t^*) and has the same interpretation as the third term. It measures the increase of value from perceptions in the transitionary dynamics.

Note that the second term is generated regardless of the chosen solution. This can be thought as the social value created from the network as users find out about the product by themselves. If the highest average cycle λ is positive, this has a positive value, but nothing prevents from λ to be negative. This obviously depends on perceptions of the most influential users.

The third statement describes the profits that the company can capture. First, the component of income implying λ disappears as the advertiser gets it for "free" if the firm performs no campaign. Second, the effortless campaign generates the vector v^* at every period and thus, it yields no income through the value of perceptions. Hence, in the periodic regime, the gain through perceptions remains as in income. Finally, since costs are subtracted and equal to half of the income that depends on signals, only the first half remains.

In the example of Figure 1(b), if $p(3\beta + 5\beta^2(1-\beta)^{-1} > 100$, the highest concentration Nash equilibrium is given by $\tilde{s}_0 = \left[p, p + p\beta, p, p, p + 4p\beta + 5p\beta(1-\beta)^{-1}\right]'$. Since $\lambda = 0$, the social value is also 0. Given the solution $\tilde{s}_0$, the network enters the periodic regime with the state associated with the eigenvector $v^* = [25, 25, 25, 26]$ after the first period. So the transient time is equal to one period ($t^* = 1$). If the vector $\mathbf{0}$ is seeded to the network, the system also enters the eigenstate v^* ($[0, 0, 0, 1]'$ is a linear combination of v^*) after one period. One can then calculate profits for the highest concentration solution:

$$\Pi(s^*, p) = \underbrace{\frac{3}{2}p^2 + \frac{1}{2}(p + p\beta)^2 + \frac{1}{2}\left(p + 4p\beta + \frac{5p\beta^2}{(1-\beta)}\right)^2}_{\text{Value of the N. eq.}}$$
$$\cdots + \underbrace{\left(-400 + \frac{\beta}{1-\beta}\right)}_{\text{Value of Perceptions}}. \qquad \text{(Highest Concentration Profits)}$$

2.6 A Complete Ordering of the Nash Equilibria

The next proposition states that the high concentration, or similarly, the equilibrium with the highest variance in the seed is profit maximizing:

Proposition 3. *Assume $G(V, E)$ admits at least one constrained star and define the variance on $\tilde{s} \in \tilde{S}$ by:*

$$var(\tilde{s}) \equiv var(\tilde{s}_0) = \frac{1}{|V|}\sum_i \left(\tilde{s}_{0,i} - \frac{p}{1-\beta}\right)^2.$$

Then, s^* maximizes the variance on $\tilde{S}$.

Proof. See Appendix A.4 $\square$

This proposition establishes a complete order on all elements in $\tilde{S}$. If one candidate solution has a smaller variance than another, then it cannot be an optimal solution. In other words, the company has an incentive to generate the highest variance. Note that a high variance is equivalent to giving most of the seed to a single user.

The fact that variance can be used as an exact metric of optimality pertains to the specificity of quadratic costs. However, the idea that the company has incentives to concentrate the value of the seed amongst users pertains to reaction function of individuals.

So this result might change in its form if another cost function is chosen, but is robust in nature.[5]

The intuition for this last result can be made quite clear. A higher signals increases the chances that it will be shared, and thus replicated. Given these choices, the firm has an interest in focusing on a few selected users to increase the value of the signal, which in turn decreases the value of other signals (since it must have a fixed average value). This increases the dispersion of signals across users[6] and thus the variance.

2.7 Long-Run Structure of Information Shared

This section explores what happens to the social network structure once the network has entered a periodic regime. In short, it becomes a constrained star network. The network is constrained because not all users can connect to the hub, as imposed initially by the topology of $G(V, E)$. So the periphery consists of nodes that are connected to the hub through a chain of users (see Figure 3). These paths are characterized, in turn, by the sequence of perceptions that maximizes the average increase of the signals over the length of the path.

[Figure 3 about here.]

The star network structure is an efficient configuration in a vast number of circumstances (see [14] ch. 6.3 for a description). Intuitively, this is the most efficient configuration of the network to maximize the increase of valuation from the center to the edges. Since not all nodes are connected to the center of the star, the network can only achieve a constrained version of it.

Proposition 4. *Let $G(V, E)$ admit at least one constrained star topology. Then after the transient time t^*, users on the network generate at least one constrained star network structure where:*

1. *All the hubs are a critical cycle;*
2. *All users that are not in the hub are connected through a path that maximizes the signal increase between the hub and the user.*

Proof. See Appendix A.5 □

This means that even if users are not aware of the whole network structure, a pattern of recommendations that maximizes the valuation of the signal emerges. There is a cycle

[5] For a strictly increasing cost function $c(s_0)$, the first order condition leads to $p\frac{|V|}{1-\beta} = \sum_i c'(s_{0i})$, so the interpretation is similar up to a positive transformation.

[6] In this simple argument, one can also see the roots of an algorithm to find the optimal solution.

of influential users at the center, and other users attach themselves to this cycle in the way that maximizes the increase of the signal. Such structure reproduces the structure found in the paper by Leskovec, Adamic and Huberman [16]. The paper finds that cascades of recommendations are led by a small group of active users (the hub) to a chain of users (the constrained periphery).

3 An Algorithm to Find the Solution

With small networks, guessing all the possible equilibria might be the fastest strategy. In large scale networks, such task would be insurmountable, however, and an algorithm is required. In this section, I present the key elements for such an algorithm. The key contribution is to show that what is known to be the "greedy" algorithm leads to an exact solution. This algorithm is known to work in polynomial time.

Recall that the average value of the seed must equal $\frac{p}{1-\beta}$. This means that the total amount of seed to distribute amongst users is $\frac{p|V|}{1-\beta}$. I call this the **budget** the algorithm has to spend on users. Because concentrating the budget to one user is the optimal solution, the greedy algorithm finds the user who can sustain the highest share of the budget given that it must be part of a Nash equilibrium. Then, it checks if the remaining of the budget can be given to a second user, and so on, until all the budget is distributed.

Below is a more elaborate description of these steps. I use the notation $v^{(i)}$ to designate the user with the ***i*th position** in the distribution of the seed. For instance, $v^{(1)}$ is the user in first position, or the user with the most valuable information.

1. Initialize the current budget to $p\frac{|V|}{1-\beta}$, the upper bound to infinity (∞) and the set of treated users to the empty set.
2. While the set of treated users has fewer than $|V|$ elements, consider an iteration i:
   (a) Initialize the seed in the following fashion. For all treated users, set their value found in the previous iterations. For untreated users, pick one node and set the seed for this particular node to the value of the minimal value between the current budget and the upper bound and set to zero all other untreated users. Compute the closest sustainable Nash equilibrium.
   (b) Repeat for all untreated users and keep the resulting solution that has the highest variance. This node is $v^{(i)}$.
   (c) Update the current budget by subtracting the value allocated to $v^{(i)}$. Update the upper bound to the maximal value such that $v^{(1)}, v^{(2)}, \ldots, v^{(i)}$ remains unchanged when seeding an untreated user.
   (d) Add user $v^{(i)}$ to the set of treated users.

A formal description of the algorithm can be found in Appendix C. The algorithm exploits two properties of the model. First, concentration is good, so it seeks to put as much of the seed as possible in a user at every step given that the average value must be $\frac{p}{1-\beta}$. Second, it exploits the fact that the solution must be a Nash equilibrium. A Nash equilibrium can be found through the application of Brouwer's fixed point theorem through the following formula:

$$\tilde{s}_{i,0} = \sum_{t=0}^{t^*} \beta^t k(i,t,s_0) + \frac{\beta^{t*+1}}{1-\beta^T} \sum_{u=0}^{T-1} \beta^u k(i,u,s_0), \tag{2}$$

where $k(i,t,s_0)$ returns the number of times a signal i is selected at time t (the vector of k_{jt}s).

I show in Appendix A.6 that if there exists a mean preserving spread of an existing Nash equilibrium, it must be that value is concentrated further in some users. This establishes a complete lattice on users and one can separate the problem in finding the most influential user $v^{(1)}$ independently of other solutions. Then, one can find the second most influential user $v^{(2)}$ given the value found for $v^{(1)}$ is part of the solution and so on. The search is thus greatly reduced as there is a lexicographic value of Nash equilibria. This property is what makes the greedy algorithm find an exact solution.

In order to prove that the steps described above are indeed an algorithm, I need to define the two key functions that I described informally above. I start by defining the function that builds a Nash equilibrium.

Definition 5 (Choice Function and Choice Value)**.** *Let $G(V,E)$ admit at least one constrained star with a transient time t^* and a period T. Define the convex and compact set*

$$C = \left\{ s_0 : s_0 \in \mathbb{R}^{|V|}, 1 \leq s_{i,0} \leq \frac{|V|}{1-\beta} \ \forall i \right\},$$

the choice function $k : C \times \mathbb{N} \to \mathbb{N}^{|V|}$ of G is defined by $k(i,t,s_0) \equiv k_{it}$ and the Choice Value function $h : C \to C$ by equation 2.

This allows it to be shown that 2 has a fixed point:

Proposition 5 (Fixed Point of Equation 2)**.** *Under the same conditions as in the last definition, the function h:*

1. *Admits a fixed point $\tilde{s}_0$;*

2. *Is a contraction.*

Proof. See Appendix A.7 □

The intuition that h is a contraction can be seen in Figure 4. Such a Figure depicts a user decision at some given time. This decision is visualized in the numbers $k_{i,t}$. On the horizontal axis, there is the current guess made by the company about the number of selections $k_{i,t}$. These guess form the actual signals $s_{i,0}$ and given this signal, a user $m \in V$ must decide which signal to choose. The solid line represents the user decision rule $(\max(s_{i,t} + b_{mi}, s_{j,t} + b_{mj}))$ and the vertical axis represents the actual choice $k'_{i,t}$ of the user given the current signals.

If the current $k_{i,t}$ is too low, other signals are higher and it does not affect the decision, that is, the flatline portion of the solid line. However, if $k_{i,t}$ is high enough, the signal is selected and leads to an increase of $k_{i,t}$. Thus, the kink of the solid line represents the threshold point at which the user change his mind.

In Panel 4(a), the current guess $k_{i,t}$ is too high since the it does not induce the user to deliver the selection guessed in the seed. Hence, the signal must be adjusted downward to satisfy first order conditions. This is repeated until k'_{jt} equals the selection, the point where the 45 degree line meets the solid line.

Conversely, Panel 4(b) shows what happens when the signal is too low. In this case, the user i is more influential than what the current seed assumes and, thus, the signal must be increased. If the signal is high enough the user making the decision becomes the unique source and all his or her friends listen to the user. In this case, the value $k_{i,t}$ equals k_{max} in the picture and the fixed point is in the top corner of the picture.

[Figures 4(a) and 4(b) about here]

In the next proposition, I show that the code stops in finite time. I also show that the execution time is of an order of $|V|^5(t^* + T)$.

Proposition 6 (Execution Time)**.** *Consider the algorithm described in the previous pages (formally described in Appendix C) on a graph $G(V, E)$. Let t^*, T be respectively the transient time and the period. Then:*

1. *Its execution time is finite;*

2. *Its execution time is at most of $O(|V|^5(t^* + T))$ for an exact solution.*

Proof. See Appendix A.8. □

4 Applications

In this section, I show how the model can answer some practical questions. I start by showing how Katz measure of influence can be erroneous when users choose their sources of information. Second, I discuss the model in the context of a partially known network structure. Third, I show how the model can be ported to an empirical framework through the discrete choice model. Finally, I show how β influences the company's optimal strategy.

4.1 Another Approach to Measure Influence

Investment in users for a word-of-mouth campaign depends on how influential each user is. If a user gets the ears of others, he will be targeted with more effort. This model shows that such analysis should be based on at least two dimensions: the structure of the network and decisions made by users based on such structure. If one dimension is missing, the measure of influence might be misleading.

A common measure of nodes on a network is **Katz prestige**. Denote A the adjacency matrix of a network,$V^{-1}\Lambda V$ its eigenvector/eigenvalue decomposition and $\imath$ the unit vector. Then, Katz prestige is defined by:

$$\begin{aligned}\rho &\equiv (I - \alpha A)^{-1} - I)\imath, \qquad \alpha \in [0,1), \\ &= \sum_{t=1}^{\infty} \alpha^t A^t \imath, \\ &= V^{-1}\left(\sum_{t=1}^{\infty} \alpha^t \Lambda^t\right) V\imath.\end{aligned}$$

Katz prestige measures how nodes are influential by amplifying signals in time (through the discounted value of eigenvalues) and how such amplification is dispersed across nodes (through eigenvectors).

This indicator assumes that all signals in the neighbourhood of a user are shared at each period. The vector $A^t\imath$ counts the number of paths of length t passing through a node; in other words, the maximal number of times a signal can be chosen each period. As such, it gives too much influence to nodes that are in multiple sets of sources.

If the discrepancy is too large, this leads to an inaccurate measure of influence. To illustrate, the Katz prestige of users in Figure 1(b) with $\alpha = \beta = 0.9$ is given by the vector $\phi \approx [0.32, 0.32, -0.64, 0.32, 0.96]'$, so it suggests that user five is the most influential node. However, for this given value of β, the high concentration equilibrium is not sustainable. Hence, the optimal solution is to seed around the periphery of the network with the low concentration equilibrium. In this equilibrium, user five is never selected by any other users!

If one uses the seed normalized by the total sum of the seed $(|V|p(1-\beta)^{-1})$ as an alternate measure of influence, one gets much more accurate description. In the case of the low concentration equilibrium, such a measure of influence yields a vector $[0.24, 0.25, 0.23, 0.24, 0.02]'$. In the case of the high concentration equilibrium, the measure yields a vector of influence $[0.02, 0.04, 0.02, 0.02, 0.90]$. In that case, user five is indeed influential as Katz prestige measure suggests.

4.2 Word of Mouth With Incomplete Information

I assumed throughout the paper that the network structure is completely known to the firm. Here, I relax this hypothesis and assume that the network can take various topologies $G_q(V,E)$ indexed by q. I also assume there is a probability distribution $f(q)$ on each of these topologies, and that a realization q^* happens after the firm has set-up its seed. The problem that the firm faces is thus:

$$\begin{aligned}
&\max_{s_0} \mathbb{E}\left[I(s_0,p) - I(\mathbf{0},p)\right] - \sum_{i\in V} \frac{s_0^2}{2}, \\
&\text{s.t. } s_{t+1} = G_{q^*}(s_t)\ \forall t
\end{aligned}$$

There are two interpretations to this problem. It can be either that the firm does not have a complete knowledge of the network structure, or that some users "disconnect" from the network. I adopt this second interpretation in the discussion below.

It can then be shown that the class of optimal solutions will satisfy:

$$\tilde{s}_{j0} = p \sum_t \beta^t \mathbb{E}\left[k_{jtq}\right],$$

where k_{jtq} is the number of times the original signal s_{j0} is chosen at time t in network $G_q(V,E)$. Such a solution introduces a new trade-off between certainty and influence. The firm still seeks to concentrate signals over the most influential agents, but these agents might change according to the network realization. The trade-off is thus weighted by each realization. Hence, introducing uncertainty on the network structure can only lead to a reduction of the seed concentration. This is intuitive, as the company is less certain of who is influential, it must cover possible cases by distributing more of the initial information. The algorithm presented in section 3 still works, but most now be executed on all realizations of the network. Hence, the total execution time now depends on the number of possible realizations.

To understand how such trade-off occurs, it is useful to distinguish some baseline cases. Each of these cases depends on the importance of the users on the network. So I discuss whether if a single user disconnects from the network with some given probability $f(q)$. Otherwise, that user remains connected with probability $1 - f(q)$. In such a case, it is convenient to rewrite the previous equation as $\tilde{s}_{j0q} = \tilde{s}_{j0} + pf(q)\sum_t \beta^t \Delta k_{jtq}$ to emphasize how the number of selections on seed j changes when the user disconnects. Because there is one less user, Δk_{jtq} is either 0 or -1, depending on if the seed $j0$ has been selected or not.

First, there is the loss due to the user's disconnection. Since the user becomes disconnected from the main source of exchanges all the associated value is lost. So in the long-run, there is a loss of $\frac{-1p\beta^{t^*+u}}{1-\beta^T}$ for at least one $u \in \{1,\ldots,T\}$.

But the most important loss occurs if the user is an influential agent on the hub of the constrained star. If that constrained star is based on the only critical cycle in G_q, users must rely on the "critical cycles of second order", that is, agents with the second highest average increase that lie on a cycle. If this second cycle exists, there is a reorganization of the long run structure of source selection. If it does not exist, then all the long run value of signal selection is lost. This means that the trade-off of the company becomes much higher. As the influential users are no longer the same users, it must decrease further the concentration of the seed.

4.3 An Example of an Empirical Application

To demonstrate the empirical potential of the model, I show how it can be used to estimate the value of perceptions between each users and thus, optimize signal selection. I do so on a constructed example based on the network shown in Figure 5. The researcher observes the structure of the network, the characteristics of the signal and the selection of sources, but does not see the biases.

So, imagine that information shared on a network is given by a vector of characteristics $x \in \mathbb{R}^k$ and that the weight given to each element is given by a vector $\beta \in \mathbb{R}^k$. The value of the signal is then given by $x\beta + b_{ij} + \epsilon_{ij}$, where ϵ_{ij} $N(0, \sigma^2)$ are idiosyncratic variations of valuation of signals. For simplicity, I assume that signals shared on the network are of an equal value.[7] Hence, a source j is selected if $\bar{b}_{ij} + \epsilon_{ij} > \bar{b}_{ik} + \epsilon_{ik}$. Of course, the biases can be identified only up to a scalar as what matters is the order. In particular, if there is no choice, there is no possible identification for biases (but the point becomes moot).

I generate a dataset of 50 periods where I record the selection of each source and then proceed to a Bayesian estimation of the perceptions. I start with a uniform prior distribution over the possible perceptions. Such perceptions are restricted to $\{0, 1, 2\}$. Given this estimation technique, the results found are displayed in Table 1.

[Table 1 about here.]

[Figure 5 about here.]

4.4 Broad or Targeted Campaigns?

Should a company be aiming for word-of-mouth campaigns that target each user equally or should it focus on one user to whom it gives all the initial information? This paper says that users should receive more importance by the company when they tend to be selected more by other users. So a users' influence is measured as the discounted value of

[7] This is done so to avoid characterizing the signal as a function of its characteristics $X\beta$ and simplify the estimation.

direct and indirect listeners they have. The profit maximizing solution is then to choose the Nash equilibria with the highest sustainable concentration of information.

This means the discount factor has an impact on the optimal solution as it influences the sustainability of each equilibrium. If β is close to zero, advertisers prefer fast, direct advertising and do not want to wait for messages to diffuse through a social network. So one can see that when $\beta = 0$, the optimal solution is the average solution for all users and the variance is zero. The average solution is the only Nash equilibrium, so it is the variance maximizing solution. Everybody is targeted equally with the signal p, regardless of the long-run influence. This is standard advertising.

However, when β approaches one, it is the opposite. The long-run steady state of signal selection is given an almost infinite value. So the relative importance of the signals reaching the steady state becomes infinite. Relatively speaking, all the initial signalling resources are focused on one user and the social network lets users build the spin. This is so because a patient company profits from the work of the network to diffuse information. It can thus focus on signals of greater value.

5 Conclusion

In this paper, I characterize the optimal solution to a word-of-mouth campaign over a social network. I do so in an environment where a company can seed information to users who choose afterwards which information they share with their neighbours. This paper thus takes the standpoint that which information is shared on a social network is a deliberate choice made by its user.

The profit maximizing solution is a unique Nash equilibrium. I show that amongst all sustainable Nash equilibria, the optimal one maximizes concentration amongst a small set of users. Such concentration of the signal to a small group of users is good because the network replicates the information free of charge. By employing such a strategy, the firm thus exploits the word-of-mouth property optimally.

I also show that after a transition period, the word-of-mouth dynamics enters a periodic regime. Users on the network listen to the same sources over and over. Their uncoordinated actions are such that the social network becomes a cascade of recommendations similar to the star network, namely a constrained star structure. The most influential users, the one with the highest perceptions, are at the center of this star of influence.

I present an algorithm that can compute the optimal solution in polynomial time on any network topology. Hence the model provides a tractable solution on large scale networks.

Finally, I present an application. I show that the model can be used to estimate a discrete choice model of signal selection and show how this allows the targeting of influential agents.

A Various Proofs

A.1 Proposition and Proof of Lemma 1

Definition 6 (Path of selection)**.** *Consider an initial signal of information s_{j0}. A path $\{e_{ij}, e_{ki}, \dots, e_{lm}, \dots, e_{no}\}$ is a **path of selection** of length T between user j to n if:*

1. *For the first position in the path, the e_{ij} is such that user i selects source j in time $t = 1$;*

2. *For position t in the path, the edge e_{lm} represents the choice of source m by user l in time t;*

3. *For position T in the path, the edge e_{no} represents the choice of source o by user n in time T;*

Lemma 1 (Dependance on the Original Signal)**.** *Let $s_{t+1} = G(s_t)$ be the law of motion of the users. Consider a path ϕ of selection from j to i in time t and its associated sum of perceptions over such path $c_{j,t} \equiv \sum_{ij:e_{ij}\in\phi} b_{ij}$. Then, for all i, $s_{i,t+1} = s_{j0} + c_{j,t}$ for some $j \in \{0, 1, 2, \dots, |V|\}$.*

Proof. This is a simple proof by recursion. At time $t = 0$, the signal $s_{i,0}$ is the signal itself ($i = j$) and $c_{i,0} = 0$. At time $t = 1$, the signal $s_{i,1}$ has a value given by some $s_{j^*,0} + b_{ij^*} = \max_{k\in\eta_i}(s_{k,0} + b_{jk})$ so $c_{i,1} = b_{ij^*}$. Now, for any given time, $s_{i,t}$ is equal to $s_{j^*,t-1} + b_{ij^*} = \arg\max k \in \eta_j(s_{k,t} + b_{jk})$ which, by recursion, can be written as $s_{j^*,0} + b_{ij^*} + c_{j,t-1}$ for some i. The conclusion follows. □

A.2 Proof of Proposition 1

The first statement is simply a constraint of the company's problem. Now, for any s^*_{j0} there are some k_{jt} at each period that counts the number of times it has been selected by users. Hence, the total value of such signal is given by:

$$p \sum_{t=0}^{\infty} k_{jt}\beta^t(s^*_{j0} + c_{j,t}).$$

First-order conditions then commands that :

$$0 = \frac{\partial \Pi(s,p)}{\partial s_{j0}}$$
$$\Rightarrow s_{j0} = p \sum_{t=0}^{\infty} \beta^t k_{jt}.$$

Hence, an optimal solution must satisfy these two conditions together. The last item stems from the fact that $\sum_{j=1}^{|V|} k_{jt} = |V|$ at any period. If there is at least one constrained star network, users can make at least one choice at any period. Hence, by the first order condition:

$$\begin{aligned}\sum_{j \in V} s_{j0} &= p \sum_{t=0}^{\infty} \beta^t \sum_{j \in V} k_{jt}, \\ &= |V| \frac{p}{1-\beta},\end{aligned}$$

from which the conclusion follows.

A.3 Proof of Proposition 2

1. After t^*, the network enters a periodic regime of period T. Hence the choices made over s^*_{j0} repeat themselves every T iterations. Hence, the optimal solution can be written as:

$$\begin{aligned}s^*_{j0} &= p \sum_{t=0}^{t^*} \beta^t k_{jt} + p\beta^{t^*} \sum_{t=1}^{\infty} \left[\beta^{tT+1} k_{j1} + \beta^{tT+2} k_{j2} + \cdots + \beta^{(t+1)T-1}\right], \\ &= p \sum_{t=0}^{t^*} \beta^t k_{jt} + p \frac{\beta^{t^*}}{1-\beta^T} \sum_{u=0}^{T-1} \beta^u k_{ju},\end{aligned}$$

   where k_{ju} is the number of times s^*_{j0} is selected in period $u \in [0, 1, \ldots, T-1]$.

2. Consider the law of motion G' :

$$s_{t+1} = G(s_t) - \begin{bmatrix} \lambda \\ \lambda \\ \vdots \\ \lambda \end{bmatrix}$$

   As the relative value of each signal has not changed, s^* is also a solution to G'. The only difference is that each step, a value of λ is substracted. So for a particular user j, the law of motion G generates:

$$s^*_{j0}, s^*_{j1} + \lambda, s^*_{j2} + 2\lambda, \ldots, s^*_{jn} + n\lambda, \ldots$$

Hence, for an user j, the discounted value of all terms involving λ can be written as:

$$\begin{aligned}I(\lambda) &= p\beta\lambda \sum_{t=0}^{\infty} \beta^t t \\ &= p\beta\lambda \left(\frac{\beta}{(1-\beta)^2}\right) \\ &= p\lambda \frac{\beta^2}{(1-\beta)^2}\end{aligned}$$

So summing over all nodes yields a value of :

$$p\lambda \frac{\beta^2 |V|}{(1-\beta)^2},$$

which is the second term of the expression. The other segments of the expression simply decompose the value of G' after and before t^*.

3. Notice first that by definition, $G^t(\mathbf{0}) = v^*$ for $t > t^*$. Using the expression found in the previous point, one can then find, the value of $I(\mathbf{0}, p)$. Profits are then just a substraction of this term and the costs of production.

A.4 Proof of proposition 3

Let $\tilde{s}, \tilde{s}' \in \tilde{S}$ be two candidate solutions. Since both solutions at time t_0 have a mean of $\frac{p}{(1-\beta)}$, one is a mean preserving spread of the other. So assume without loss of generality that :

$$\begin{aligned}\tilde{s}'_0 &= \tilde{s}_0 + \Delta\tilde{s}_0, \\ \text{with } 0 &= \sum_i \Delta\tilde{s}_{0,i}.\end{aligned}$$

Consider now an agent m who chooses source j at time t under $\tilde{s}'$. This means that:

$$\begin{aligned}&\Delta\tilde{s}_{j0} + \tilde{s}_{j0} + c_{j,t} > \tilde{s}_{i,0} + \Delta\tilde{s}_{i,0} + c_{i,t} \ \forall i \\ \Leftrightarrow\ &(\Delta\tilde{s}_{j0} - \Delta\tilde{s}_{i,0}) + (\tilde{s}_{j0} - \tilde{s}_{i,0}) + (c_{j,t} - c_{i,t}) > 0 \ \forall i.\end{aligned} \tag{7}$$

Consider also the difference in profits due to variance:

$$\frac{1}{2}\sum_{i\in V}\left(\left[\tilde{s}_{i,0} + \Delta\tilde{s}_{i,0} - \frac{p}{1-\beta}\right]^2 - \left[\tilde{s}_{i,0} - \frac{p}{1-\beta}\right]^2\right) = \frac{1}{2}\sum_{i\in V}\Delta\tilde{s}_{i,0}\left(2\tilde{s}_{i,0} + \Delta\tilde{s}_{i,0}\right) > 0$$

This implies that :

$$
\begin{aligned}
0 &< \sum_{i\in V}(\Delta\tilde{s}_{i,0})^2 + \sum_{i\in V}\Delta\tilde{s}_{i,0}\tilde{s}_{i,0}, \\
&= \sum_{i\in V}\Delta\tilde{s}_{i,0}(\Delta\tilde{s}_{i,0}+\tilde{s}_{i,0}), \\
&= \sum_{i\in V}\Delta\tilde{s}_{i,0}\tilde{s}'_{i,0}.
\end{aligned}
$$

Now, consider the total change in profits:

$$
\begin{aligned}
\Delta\Pi(\tilde{s},\tilde{s}',p) &= \frac{1}{2}\sum_{i\in V}\Delta\tilde{s}_{i,0}\left(2\tilde{s}_{i,0}+\Delta\tilde{s}_{i,0}\right) + \sum_{t=0}^{\infty}\beta^t\sum_{m\in V}\sum_{i\in V}\mathbb{1}_{mit}\left[\tilde{s}_{j0}+c_{j,t}+\Delta\tilde{s}_{j0}-\tilde{s}_{i,0}-c_{i,t}\right], \\
&= \frac{1}{2}\sum_{i\in V}\Delta\tilde{s}_{i,0}\left(2\tilde{s}_{i,0}+\Delta\tilde{s}_{i,0}\right) + \sum_{t=0}^{\infty}\beta^t\sum_{i\in V}\tilde{k}_{it}\left[\tilde{s}_{j0}+c_{j,t}+\Delta\tilde{s}_{j0}-\tilde{s}_{i,0}-c_{i,t}\right], \\
&= \frac{1}{2}\sum_{i\in V}\Delta\tilde{s}_{i,0}\left(2\tilde{s}_{i,0}+\Delta\tilde{s}_{i,0}\right) + \sum_{t=0}^{\infty}\beta^t\sum_{i\in V}\tilde{k}_{it}\left[\tilde{s}_{j0}+c_{j,t}+\Delta\tilde{s}_{j0}-\tilde{s}_{i,0}-c_{i,t}\right]\dots \\
&\quad \dots + \sum_{i\in V}\Delta\tilde{s}_{i,0}\tilde{s}_{i,0} - \sum_{i\in V}\Delta\tilde{s}_{i,0}\tilde{s}_{i,0}.
\end{aligned}
$$

By construction, we have :

$$
\begin{aligned}
\sum_{i\in V}\Delta\tilde{s}_{i,0}\tilde{s}_{i,0} &= \sum_{i\in V}\Delta\tilde{s}_{i,0}\left(\sum_{t=0}^{\infty}\beta^t\tilde{k}_{i,t}\right), \\
&= \sum_{t=0}^{\infty}\beta^t\sum_{i\in V}\tilde{k}_{i,t}\Delta\tilde{s}_{i,0}.
\end{aligned}
$$

Hence, we can combine the fourth term into the second term and the third term into the first term:

$$
\Delta\Pi(\tilde{s},\tilde{s}',p) = \frac{1}{2}\sum_{i\in V}\Delta\tilde{s}_{i,0}\left(3\tilde{s}_{i,0}+\Delta\tilde{s}_{i,0}\right) + \sum_{t=0}^{\infty}\beta^t\sum_{i\in V}\tilde{k}_{it}\underbrace{\left[\tilde{s}_{j0}+c_{j,t}+\Delta\tilde{s}_{j0}-\Delta\tilde{s}_{i,0}-\tilde{s}_{i,0}-c_{i,t}\right]}_{>0}.
$$

By equation 7, the last term is positive. Now, for the first term, consider that:

$$
\begin{aligned}
\frac{1}{2}\sum_{i\in V}\Delta\tilde{s}_{i,0}\left(3\tilde{s}_{i,0}+\Delta\tilde{s}_{i,0}\right) &= \frac{1}{2}\sum_{i\in V}\Delta\tilde{s}_{i,0}\left(3\tilde{s}_{i,0}+\Delta\tilde{s}_{i,0}\right) + \frac{1}{2}\sum_{i\in V}\left(\frac{3}{2}\tilde{s}_{i,0}\right)^2 - \frac{1}{2}\sum_{i\in V}\left(\frac{3}{2}\tilde{s}_{i,0}\right)^2, \\
&= \frac{1}{2}\sum_{i\in V}\left(\frac{3}{2}\tilde{s}_{i,0}+\Delta\tilde{s}_{i,0}-\frac{3}{2}\frac{p}{1-\beta}\right)^2 - \frac{1}{2}\sum_{i\in V}\left(\frac{3}{2}\tilde{s}_{i,0}-\frac{3}{2}\frac{p}{1-\beta}\right)^2,
\end{aligned}
$$

which is positive.

Since all terms in the difference in profits are increasing with a mean preserving spread, the variance is a sufficient measure for maximizing profits.

A.5 Proof of Proposition 4

1. Assume that a user that is not on a critical cycle is in the hub. By definition, this means that there exits a directed cycle starting at i. But since i is not in a critical cycle, this means that the increase λ' over the cycle is smaller than the increase on a critical cycle. Thus, for any s^*_k in the periodic regime, $s^*_k + T\lambda' < s^*_k + T\lambda$. Thus, the agent does not select this cycle as a source. By the same token, it does not select any other possible cycle that is not a critical cycle. Hence, the user cannot be in the hub. Hence, only users in critical cycles can be in the hub.

2. Assume that a path does not maximize the increase of the signal. This means that there exists another path from the hub with an higher increase of the signal. So this means that for the same signal, one has chosen $s^*_{k,0} + c_{kt} < s^*_{k,0} + c'_{kt}$, which contradicts that users choose the highest utility signal.

A.6 Proof of Lemma 2

Lemma 2. *Let $\tilde{s}, \tilde{s}'$ be two Nash Equilibria on $G(V, E)$ and assume that $\tilde{s}'_0$ is a mean-preserving spread of $\tilde{s}_0$ so that :*

$$\tilde{s}'_{i,0} = \tilde{s}_{i,0} + \Delta s_{i,0},$$
$$\textit{with } 0 = \sum_{i \in V} \Delta s_{i,0}.$$

Then, the covariance between the solutions is positive:

$$\sum_{i \in V} \Delta \tilde{s}_{i,0} \tilde{s}_{i,0} > 0.$$

Proof. See Appendix A.6. ◻

This is a simple corollary of the increase of variance of proposition 3. Since we have:

$$\frac{1}{2} \sum_{i \in V} (\tilde{s}_{i,0} + \Delta \tilde{s}_{i,0})^2 - \frac{1}{2} \sum_{i \in V} (\tilde{s}_{i,0})^2 > 0,$$
$$\Rightarrow \underbrace{\frac{1}{2} \sum_{i \in V} \Delta \tilde{s}_{i,0} \tilde{s}_{i,0}}_{\text{Covariance}} + \underbrace{\frac{1}{2} \sum_{i \in V} \Delta \tilde{s}^2_{i,0}}_{>0}.$$

Hence, for two identical increases $\Delta s_{i,0}$, the one allocated to an already high solution will increase profits more.

A.7 Proof of Proposition 5

For the first element of the proposition, it is sufficient to notice that h is a mapping a convex and compact set to itself. By Brouwer's Fixed Point Theorem, h admits a fixed point.

For the second element of the proposition, denote s_0' the result of $h(s_0)$ for some s_0 and consider a change $\Delta k_{it}'$ from user m at time t. Assume this user changes from source j to source i. For $s_{i,0}'$ to increase, it must be that:

$$\begin{aligned}
s_{i,0} + \beta^t \Delta k_{i,t}' + c_{i,t-1} + b_{mi} &> s_{j0} - \beta^t \Delta k_{i,t}' + c_{j,t-1} + b_{mj} > 0, \\
\Rightarrow 2\beta^t \Delta k_{i,t}' &> s_{j0} - s_{i,0} + c_{j,t-1} - c_{i,t-1} + b_{mj} - b_{mi} > 0,
\end{aligned}$$

for some user m. At the same time, we must have that $\eta_m - k_{it} \geq 2\beta^t \Delta k_{i,t}'$ for all values of t. So this leads to the following constraint:

$$\eta_m - k_{it} > 2\beta^t \Delta k_{i,t}' > s_{j0} - s_{i,0} + c_{j,t-1} - c_{i,t-1} + b_{mj} - b_{mi} > 0.$$

Hence, if $2\beta^t \Delta k_{i,t} \geq \eta_m - k_{it}$, we have that Δk_{it} must be reduced. So any increase in k_{it} is bounded by the neighborhood of m and previous choices.

Now consider the decision of user m' who faces the change Δk_{jt} but does not change his decision from source j to source k, the second highest source. For this to hold, it must be that:

$$\begin{aligned}
s_{j0} - s_{k,0} + c_{j,t-1} - c_{k,t-1} + b_{mj} - b_{mk} &> -\beta^t \Delta k_{jt}, \\
\Rightarrow s_{j0} - s_{k,0} + c_{j,t-1} - c_{k,t-1} + b_{mj} - b_{mk} &> \beta^t \Delta k_{i,t}
\end{aligned}$$

Hence, if s_{j0} is decreasing without any change, its decrease is bounded by the value of the second highest source.

By Lemma 2, if a signal increases (decreases), it will keep increasing (decreasing). But the changes in ks are bounded for any t. Thus, they converge to a point in $\mathbb{N}$, which means that h converges as well.

A.8 Proof of Proposition 6

1. Each time the function is called, the number of nodes treated increases by at least one. Hence, the stopping condition is reached eventually. Now since h converges on $\mathbb{N}$, the number of steps required to reach convergence is finite. Hence, the algorithm stops in finite time.

2. The findValue is called at most $|V|^2$ times in the general algorithm. Such function converges in the worst case in $|V|(t^* + T)$ and each step requires $|V|^2$ computations. So the computational speed of findValue is of $O(|V|^t(t^* + T))$.

   The function findUpperBounds is called $|V|$ times in the general algorithm and requires $|V|^2(t^* + T)$ steps to compute G^t, $|V|^2(t^* + T)$ steps to compute all values and it calls the function sort $|V|$ times. A sort function takes at most $|V|^2$ steps to sort. Hence, the function algorithm requires an order of $O(|V|^5(t^* + T)$ steps to use findUpperBounds. The conclusion follows.

B A Primer on Tropical Algebras

In this appendix, I introduce various definitions of the mathematical algebra I use in the last segment of the paper, namely a tropical algebra. This algebra is useful for allowing the problem to be described in a linear space, which reveals a tractable analytic solution.

Intuitively, this algebra is almost identical to the classical high school algebra except for one property: there is no unique inverse under addition. Hence $a \oplus x = b$ has no unique solution for x. Otherwise, the properties are similar.

Definition 7 (Tropical Algebra)**.** *A tropical algebra is a semiring over a set* $\mathbb{R}_{\max} \equiv \mathbb{R} \cup \{\infty\}$ *with an addition operator and a multiplication operator defined by:*

$$a \oplus b \equiv \max\{a, b\} \;\; \forall a, b \in \mathbb{R}_{\max}$$
$$a \otimes b \equiv a + b \; \forall a, b \in \mathbb{R}_{\max}$$

With such an algebra, the multiplicative and additive identities are given by $e = 0$ and $\epsilon = -\infty$. One can easily check that some usual properties of standard algebra (and some others) are met:

$$x \oplus (y \oplus z) = (x \oplus y) \oplus z, \qquad x \otimes (y \otimes z) = (x \otimes y) \otimes z \qquad \text{(associativity)}$$
$$x \oplus y = y \oplus x \qquad x \otimes y, = y \otimes x \qquad \text{(commutativity)}$$
$$x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z) \qquad \text{(distributivity)}$$
$$x \oplus \epsilon = \epsilon \oplus x = x \qquad \text{(zero element)}$$
$$x \otimes e = e \otimes x = x \qquad \text{(unit element)}$$
$$x \otimes \epsilon = \epsilon \otimes x = \epsilon \qquad \text{(absorbing element)}$$
$$\forall x \in \mathbb{R}, \exists! y : x \otimes y = e \qquad \text{(unique multiplicative inverse)}$$
$$x \oplus x = x \qquad \text{(idempotency of addition)}$$

Unnacustomed readers might want to check that the following equations are true:

$$10^2 \equiv 10 \otimes 10 = 20 \qquad (-0.5)^2 = -1$$
$$6 = 2 \otimes 4 \oplus 5 \qquad \sqrt{-1} = -0.5$$
$$7 = 3 \otimes 2^2 \oplus 1 \otimes 2 \oplus 1 \qquad 4^5 = 5^4 = 20$$

Matrix algebra can also be defined. I describe here its relationship with a network:

Definition 8 (Adjacency Matrix)**.** *Define the weighted adjacency matrix $B \in \mathbb{R}_{\max}^{|V|\times|V|}$ by :*

$$[B_{ij}] = \begin{cases} b_{ij} & \text{if } j \in \eta_i, \\ -\infty & \text{if } j \notin \eta_i. \end{cases}$$

The adjacency matrix contains the entries of perceptions that must be added from each signal sent from the neighborhood. Users that are not in the neighborhood of user i are weighted with the neutral element $(-\infty)$ under $\oplus$ and are thus never chosen (not accessible). For instance, the matrix of the network presented in Figure 1(b) is given by :

$$B = \begin{bmatrix} -\infty & -\infty & -\infty & 0 & -100 \\ 0 & -\infty & -\infty & -\infty & -100 \\ -\infty & 0 & -\infty & -\infty & -100 \\ -\infty & -\infty & 0 & -\infty & -100 \\ -\infty & 1 & -\infty & -\infty & -\infty \end{bmatrix}.$$

Definition 9 (Irreducible Matrix)**.** *Let $B \in \mathbb{R}_{\max}^{|V|\times|V|}$, the matrix B is irreducible if it spans the full $|V|$-dimensional vector space or equivalently, if its determinant $|B|$ is not equal to ϵ.*

Definition 10 (Inner Product)**.** *Let $a, b \in \mathbb{R}_{\max}^{|V|}$ be two vectors in a tropical space. Their inner product is defined as*

$$a \cdot b = \bigoplus_{i \in V} a_i \otimes b_i.$$

With such a definition, the global behavior of the network can be modeled as a homogenous, first-difference equation:

$$s_{t+1} = B \cdot s_t$$

where B depicts the perceptions for each pair of users.

Such a matrix system then allows them to be defined in terms of eigenvalues and eigenvectors.

Definition 11 (Eigenvalues and Eigenvectors)**.** *Let $B \in \mathbb{R}_{\max}^{|V|\times|V|}$. An eigenvalue λ of such matrix is a solution to the equation*

$$B \cdot v = \lambda \otimes v.$$

A particular vector v satisfying this equation for a given λ is called an eigenvector.

The (unique) eigenvalue of a network has an interpretation that depends on the notion of cycles and critical cycles. Every possible cycle in the network has an average net-spin increase over its edges. The eigenvalue is the highest average net spin.

Definition 12 (Cycles)**.** *A simple cycle of length n in a graph $G(V, E)$ is a set of edges $\{a_{12}, a_{23}, \ldots, a_{n-1n}\} \subseteq E$ such that the first node is the last node $(1 = n)$ and that no edge is repeated.*

Definition 13 (Critical Cycles)**.** *Let C be the set of all cycles on $G(V, E)$. For any cycle $c \in C$ of length $|c|$, let*

$$\rho(c) \equiv \sum_{e_{ij} \in c} \frac{b_{ij}}{|c|}$$

be the average value of perceptions and $C(G)$ be the set of all cycles in a network $G(V, E)$. A cycle c_c is critical if

$$c_c \in \arg \max_{c \in C(G)} \rho(c).$$

Denote $\lambda \equiv \max_{c \in C(G)} \rho(c)$.

For instance, the set of edges $\{e_{12}, e_{21}\}$ is a cycle (a loop) in Figure 1(b). It is a critical cycle since the average net perception over this loop $\left(\frac{1+1}{2}\right)$ is the highest amongst all possible cycles.

The following theorem stitches all these definitions together:

Theorem 2 (Mathematical Results)**.** *Let $G(V, E)$ be an irreducible graph and let B be its adjacency matrix. Then:*

1. *B has a unique eigenvalue λ*
2. *There exist eigenvectors $B_1^*, \ldots B_p^*$ that span part of the graph space. Such vectors are the different permutations of the cycles belonging to $C_c(G)$.*
3. *Define $B^* \equiv \bigoplus_{i=0}^{t^*} (B \otimes \lambda^{-1})^{pi}$ where t^* is the transient time of the graph, then the different modes of the unique eigenvector are represented by the columns in B^* for which nodes are in at least one $C_c(G)$.*
4. *For any s_0, the system will enter a steady state after t^* periods. The steady state is characterized by $s_{t+p} = s_t - \lambda p$, and $s_t = B^* \cdot s_0$ when $t > t^*$.*

5. *The transient time is finite.*

Proof. See Bacelli, Cohen, J. Olsder and Quadrat [3]. □

Each column in B^* represents a different mode of the periodic regime after $t > t^*$, so each of them has a neat interpretation. The components v_i of the eigenvector j represents the net increase of the signal s_{j0} due to perceptions in the given mode.

One can verify that the vectors $v_1 = [0, 25, 0, 0, 1]'$, $v_2 = [0, 0, 25, 0, 26]'$, $v_3 = [0, 0, 0, 25, 1]'$ and $v_4 = [25, 0, 0, 0, 1]'$ are eigenvectors of the adjacency matrix associated to Figure 1(b):

$$B^4 \cdot v_i = v_i \ \forall i \neq j,$$
$$B \cdot v_i = v_{(i+1) \mod 4} \ \forall i,$$
$$B \cdot (v_1 \oplus v_2 \oplus v_3 \oplus v_4) = (v_1 \oplus v_2 \oplus v_3 \oplus v_4) = v^*$$
$$B^* \cdot v_i = v_i \ \forall i.$$

C A Formal Description of the Algorithm

Such an algorithm has been coded in Python and is available on demand. As it is longer than fifteen pages long, it is not included in the appendix.

Data: $G(V, E), \beta \in [0, 1)$ and $p \in \mathbb{R}^+$.
Result: $s_0^* \in \tilde{S}$.
Find λ;
Find t^* and T;
Initialize the set of nodes to treat to the entire set of nodes $v_t \leftarrow V$;
Initialize some seed for the solution: $s_0^* \leftarrow [0, 0, \ldots, 0]'$;
Initialize $maxBudget_i \leftarrow p\frac{|V|}{1-\beta}\ \forall i \in v_t$;
Initialize $upperBounds_i \leftarrow \infty\ \forall i \in v_t$;
while $v_t \neq \emptyset$ **do**
  $\tilde{s}_{\text{candidates}} \leftarrow \emptyset$;
  for $i \in v_t$ **do**
    for $j \in v_t$ **do**
      $s_{j,0}^* \leftarrow 0$;
    end
    $s_{i,0}^* \leftarrow maxBudget_i$;
    $s_{\text{candidates}} \leftarrow s_{\text{candidates}} \cup \{$`findValue`
    $(G, v_t, s_0^*, upperBounds, \beta, p, t^*, T, maxBudget))\ \}$;
  end
  $(i) \leftarrow i : i = \arg\max_{\tilde{s}_i \in s_{\text{candidates}}} I(\tilde{s}_i) - c(\tilde{s}_i)$;
  $s_0^* \leftarrow \tilde{s}_{(i)}$;
  $v_t \leftarrow v_t \cap (i)$;
  $upperBounds \leftarrow$ `findUpperBounds`(G, v_t, s_0^*, t^*, T);
  for $i \in v_t$ **do**
    $maxBudget_i \leftarrow \min(maxBudget_i - s_{(i),0}^*, upperBounds_i)$;
  end
end

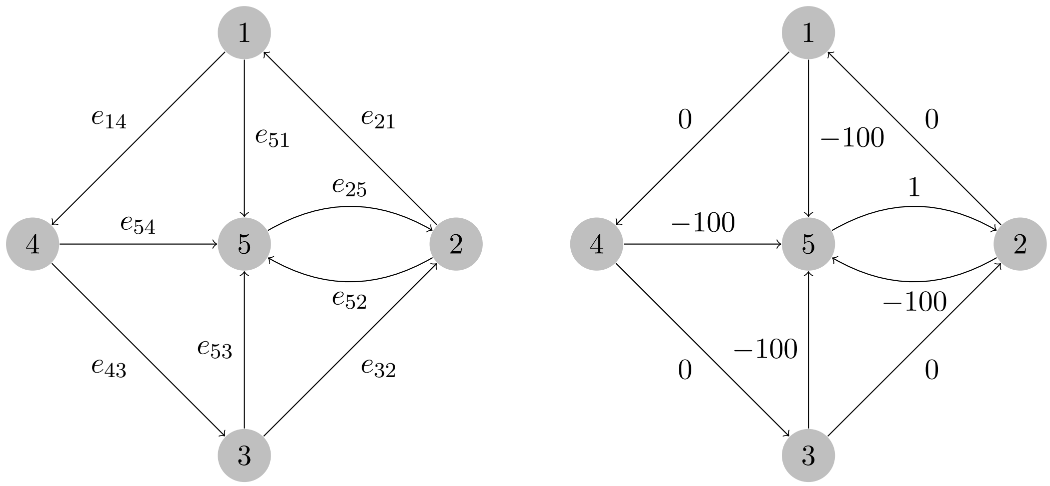
Algorithm 1: A Simple Algorithm to Find the Optimal Solution

Input: $G(V,E), v_t \subseteq V, s_0^* \in \mathbb{R}^{|V|}, upperBounds \in \mathbb{R}^{|V|}, \beta \in [0,1), p \in \mathbb{R}^+, t^* \in \mathbb{N}, T \in \mathbb{N}$ and $maxBudget \in \mathbb{R}$.
Result: $s_0^* \in \tilde{S}$.
Set $\tilde{s}_0' \leftarrow \tilde{s}_0^*$;
Set $[\tilde{s}_0']_i \leftarrow [\tilde{s}_0']_i + 1$;
while $\|\tilde{s}_0 - \tilde{s}_0'\| \geq 0$ **do** Find the closest Nash equilibrium that supports such increase:
    $\tilde{s}_0 \leftarrow \tilde{s}_0'$;
    $\tilde{s}_0' = h(\tilde{s}_0)$;
end
return $\tilde{s}_0'$;

Algorithm 2: The Function findValue Approaches s^*

Input: $G(V,E), v_t \subseteq V, s_0^* \in \mathbb{R}^{|V|}, t^* \in \mathbb{N}, T \in \mathbb{N}$.
Result: $upperBounds \in \mathbb{R}^{|V|}$.
$upperBounds_i \leftarrow \infty \ \forall i$;
for $t \in [0, \ldots, t^* + T]$ **do**
    $A \leftarrow A^{\otimes t}$;
    for $i \in V$ **do**
        $values = \emptyset$;
        for $j \in V$ **do**
            $values \leftarrow values \cup (value = A_{ij} + s^*_{j,0}, index = j)$;
        end
        $values \leftarrow \mathtt{sort}(values, \text{sort on } value, \text{descending})$;
        $chosenSource = value_0$;
        if $chosenSource.index \notin v_t$ **then**
            for $j \in v_t$ **do**
                $currentItem \leftarrow values.find(index = j)$;
                $currentValue \leftarrow currentItem.value$;
                $upperBounds_j \leftarrow$
                $\min\left(upperBounds_i, chosenSource.value - currentValue + s^*_{j0}\right)$;
            end
        end
    end
end
return $upperBounds$;

Algorithm 3: The Function findUpperBounds Ensures Stability on Past Solutions



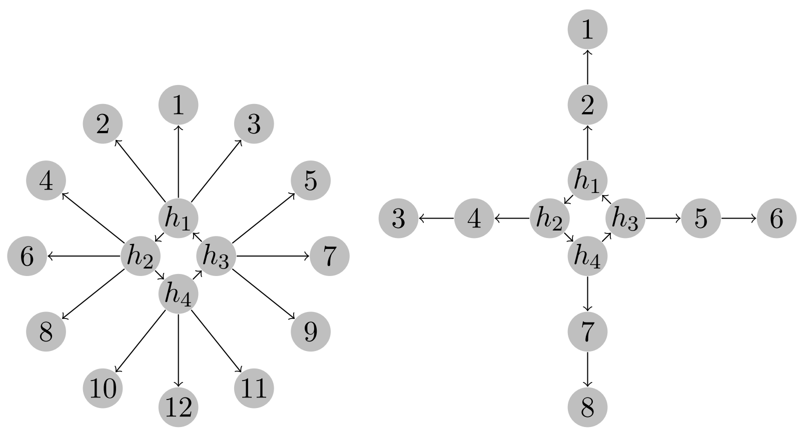
(a) **Notation for Vertices and Edges** (b) **Some Perceptions $\mathbf{b_{ij}}$ Along Edges**

Figure 1: **The Example Used Throughout This Paper**

D Figures

Figures for "Getting the Right Spin: A Theory of Value of Social Networks" Pier-André Bouchard St-Amant[8]

[8]pabsta@econ.queensu.ca.



(a) **A Star Network** (b) **A Constrained Star Network**

Figure 2: **The Star and Constrained Star Networks.**

Table 1: Estimated Probabilities of Perceptions (Simulation Based on Figure 5)

User, Source	$P(b_{ij}=0)$	$P(b_{ij}=1)$	$P(b_{ij}=2)$
1, 2	0.982	0.018	0.000
1, 5	0.000	0.316	0.684
2, 3	0.999	0.001	0.000
2, 5	0.000	0.019	0.981
3, 4	0.980	0.020	0.000
3, 5	0.000	0.342	0.658
4, 1	0.151	0.404	0.445
4, 5	0.117	0.802	0.083
5, 2	0.333	0.333	0.333

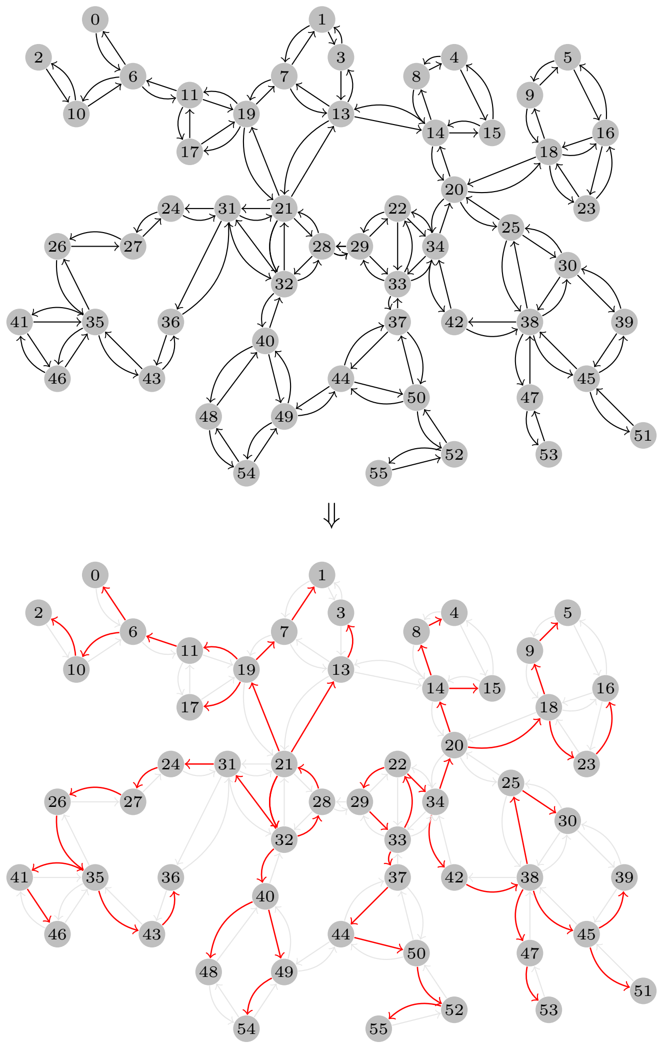
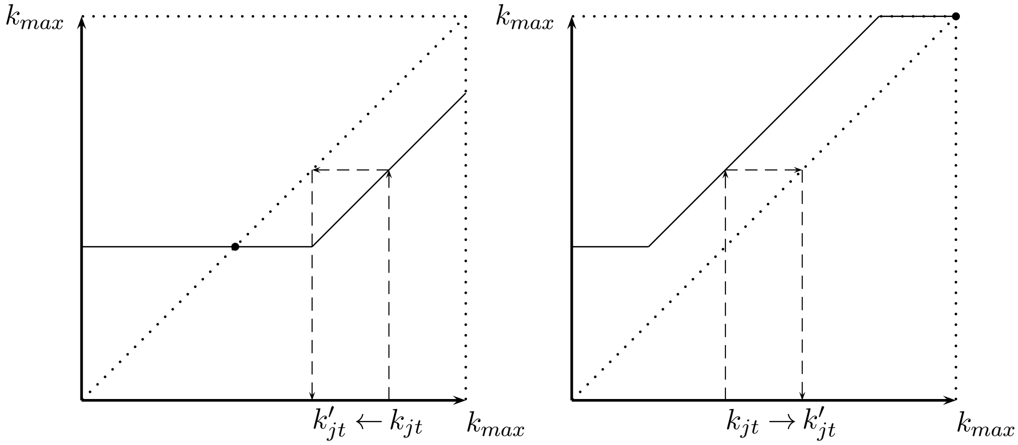


Figure 3: **The Most Complicated Networks Generate a Cascade of Recommendations**



(a) **The Signal Is Too High** (b) **The Signal Is Too Low**

Figure 4: **The Function h is a Contraction**

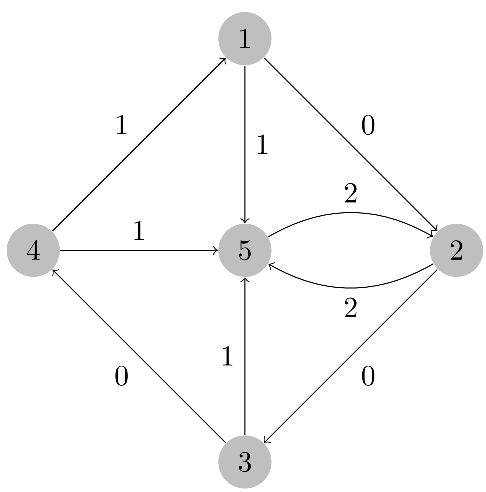


Figure 5: **An Applied Example: Bias Detection**

References

[1] Daron Acemoglu, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. The Network Origins of Large Economic Downturns. *Econometrica*, July 2012.

[2] David Arthur, Rajeev Motwani, Aneesh Sharma, and Ying Xu. Pricing Strategies For Viral Marketing On Social Networks. *Internet and Network Economics*, 5929:101–112, 2009.

[3] François Baccelli, Guy Cohen, Geert Jan Olsder, and Jean-Pierre Quadrat. *Synchronization and Linearity, An Algebra for Discrete Event Systems*, volume 45. John Wiley and Sons, 1992.

[4] Abhijit V Banerjee. A Simple Model of Herd Behavior. *The Quarterly Journal of Economics*, 107(3):797 – 817, 1992.

[5] Robert M. Bond, Christopher J. Fariss, Jason J. Jones, Adam D. I. Kramer, Cameron Marlow, Jaime E. Settle, and James H. Fowler. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295–298, September 2012.

[6] Christian Borgs, Michael Brautbar, Jennifer Chayes, and Brendan Lucier. Influence Maximization in Social Networks: Towards an Optimal Algorithmic Solution. December 2012.

[7] Arthur Campbell. Word of Mouth and Percolation in Social Networks. *American Economic Review*.

[8] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to Algorithms, Third Edition*. The MIT Press, 2009.

[9] Matthew Elliott, Benjamin Golub, and Matthew O. Jackson. Financial Networks and Contagion. *SSRN Electronic Journal*, January 2012.

[10] Johannes Fürnkranz, Tobias Scheffer, and Myra Spiliopoulou, editors. *Knowledge Discovery in Databases: PKDD 2006*, volume 4213 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[11] Andrea Galeotti and Sanjeev Goyal. Influencing the influencers: a theory of strategic diffusion. *The RAND Journal of Economics*, 40(3):509–532, 2009.

[12] Benjamin Golub and Matthew O. Jackson. Network Structure and the Speed of Learning: Measuring Homophily Based on its Consequences. *SSRN Electronic Journal*, March 2011.

[13] Morris H. De Groot. Reaching A Consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

[14] Matthew O Jackson. *Social and Economic Networks*, volume 39. Princeton University Press, 2008.

[15] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '03*, page 137, New York, New York, USA, August 2003. ACM Press.

[16] Jure Leskovec, Lada A. Adamic, and Bernardo A. Huberman. The dynamics of viral marketing. *ACM Transactions on the Web*, 1(1):5–es, May 2007.

[17] Manuel Mueller-Frank. A general framework for rational learning in social networks. *Theoretical Economics*, 8(1):1–40, January 2013.

[18] Pasquale Schiraldi and Ting Liu. New Product Launch: Herd Seeking Or Herd Preventing. *Economic Theory*, 48(1):1–22, 2011.

[19] Raj Sethuraman, Gerard J Tellis, and Richard A Briesch. How Well Does Advertising Work? Generalizations from Meta-Analysis of Brand Advertising Elasticities. *Journal of Marketing Research*, 48(3):457–471, June 2011.

[20] Michael Trusoc, Randolph E. Bucklin, and Koen Pauwels. Effects of Word-of-Mouth Versus Traditional Marketing: Findings from an Internet Social Networking Site - American Marketing Association. *Journal of Marketing*, 73(5), 2009.

[21] J.-H. Steffi Yang. Social network influence and market instability. *Journal of Mathematical Economics*, 45:257 – 276, 2009.

[22] Hema Yoganarashimhan. Impact of Social Network Structure On Content Propagation: A Study Using YouTube Data. *Quantitative Marketing and Economics*, 10(1):111–150, 2011.