Opinion dissemination and degrees of influence
Project Network Science 2021 / 2022

Alexandre
IST

Inês Correia
IST 89464
inesmargarida1618
@tecnico.ulisboa.pt

Pedro Galhardo
IST 89517
pedro.galhardo
@tecnico.ulisboa.pt

Abstract

In an increasingly connected world, studying how we're being influenced by our peers becomes more and more important. We set out to examine how influence works and its reach in networks of people. We used the models and the networks from the original paper we're reproducing, and tried different settings to see how it would influence the results. We also analyzed what could lead to the differences between our results and the original ones. Similar results were obtained and the 3-degree rule was consistent throughout experimentation, despite slight variations in the graphs.

1 Introduction

Social relations are at the core of the day-to-day of human beings. What we like, what we say and what we do isn't just affected by our own personality - instead, it's severely influenced by the people around us. And if before we already had our closest friends help shape our worldview, with conflicting opinions and habits, with the appearance of the Internet and social media, that impact on what we think has increased substantially. The average American has about 16 friends [6] - with whom they communicate, gather and schedule things. On the other hand, in 2011 the average person on Facebook had 190 friends [10]. Just from this, we can see how there's much more stimuli and different ideas interacting with us on a daily basis, much more than before. And while it's easy to understand how our direct friends influence our behavior, past studies suggest that your friends' friends and even your friends' friends' friends can weigh on how you think and see the world.

What we intended to do is recreate the models from one of these papers - Origin of Peer Influence in Social Networks - to take our own conclusions on the results and observe correlations among nodes at different distances. Hopefully, we will be able to verify the results previously obtained.

2 The origin of 3-degrees of influence

The concept was originally proposed by Nicholas A. Christakis and James H. Fowler, in their 2007 paper about obesity [2], where they observed this phenomenon over the years.

They studied 12,067 people over 32 years and took their measurements. Not only did they have massive amounts of nodes, the data was also very interesting, since they would also track obesity among family members, spouses and friends of the participants, as well as other traits, such as smoking habits or geographic distance of the peers.

At the end of the study, they managed to conclude that while direct social relations had the biggest influence on the subject's weight, that influence could be seen up to their friends' friends and their friends' friends' friends - hence the 3-degree rule. Past this point, there was no observable correlation. Besides this, the type of relationship also made a difference in regards to the amount of impact on the person.

After releasing this paper, they managed to recreate this phenomenon for smoking [3], happiness [4] and in 2009 wrote a book about it [5]. This was observed to be a general trend. They noticed that, in these several different domains, there was always this cascading (and decreasing with distance) pattern of influence among networks of friends.

3 Origin of Peer Influence in Social Networks

With the same phenomenon in mind, here it was set out to be explored from a more theoretical

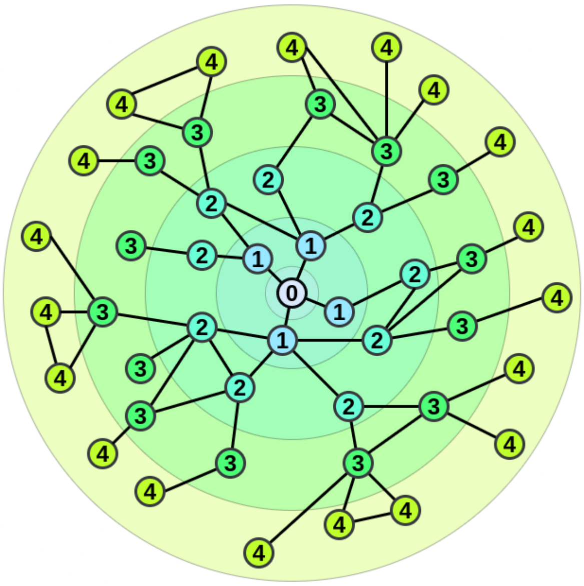

Figure 1: Network representing social relationships. In the middle, there's a node (0), representing an individual and the links are the connections to his peers. The numbers in the other nodes represent the shortest path between them and the individual. In the 3-degrees of influence theory, the influence spreads to the node at distance 3 from the center.

point-of-view - by modelling "the spread of cooperative strategies, opinions, and diseases" [7]. Across four different types of networks and three models, the authors drew their own conclusions about this theory - without relying on data - and instead, focusing on network science.

Even with all of these different approaches, the main result was the same: "Despite the importance of social networks in defining the paths and ends of the dynamical processes they support, showing how important it is to address and understand population dynamics from a complex networks perspective, the patterns of peer influence they exhibit are surprisingly independent of their structure" [7]. Although, as the authors explain, some parameters of a network can change how far an individual's influence can spread.

4 Approach

Our intention was to replicate what was done in the paper mentioned above [7].

Three models are described in the paper to simulate opinion spreading and dynamics: Evolutionary Prisoner's Dilemma; Voter Model; and SIR Model.

Furthermore, the simulation was also conducted in the four different network configurations: Homogeneous Small World; Heterogeneous Small World; Barabási and Albert; and Exponential. In these networks, the nodes represent the individuals and the links between nodes in these networks are interpreted as social relationships. All the nodes connected to a node are friends of that node (distance 1), all the nodes connected to those ones are his friend's friends (distance 2) and so on.

Each individual in the network can have two states: 1 - influenced by a belief/opinion and spreading it; and 0 - susceptible or ignorant. However, in the case of the SIR model, they can also have a third state: 2 - no longer sharing the belief with its peers (recovered).

To calculate the degree of influence a node (at a distance n) has on another node, we used a quantity called propensity [7]. It compares each one of the three Opinion Spreading models with a similar model but with a random arrangement of individuals who share the belief.

4.1 Opinion Spreading Models

Evolutionary Prisoner's Dilemma (PD). This model is often used in network science to study the dynamics of cooperation between individuals.

In the famous Prisoner's Dilemma Game, each of two players has to choose simultaneously whether they will Cooperate or Defect. For this problem, there are four combinations of outcomes possible and the respective rewards are usually represented in a 2x2 payoff matrix. If both choose to Cooperate the payoff for both is $R = 1$, and for Mutual Defection $P = 0$. However, choosing to Cooperate while the other chooses Defect, will give payoff $S = -\lambda$ and to the other player $T = 1 + \lambda$. If $\lambda > 0$ both of the players will have the temptation to defect which will lead to mutual defection, the pure Nash Equilibrium for this problem. For our implementation, we chose to use $\lambda = 1$, leaving us the following payoff matrix:

A/B	Cooperate	Defect
Cooperate	1, 1	-1,2
Defect	2,-1	0, 0

For the evolutionary version of the game, each player plays the game with his neighbors and his total payoff, or "fitness", is the sum over all the interactions. Now, when choosing the strategy (Cooperate or Defect), the player will measure his fitness in comparison with his neighbors' fitness.

Applying this to the context of our study, an individual A, with fitness f_A, is influenced by the opinion of another individual B, with fitness f_B, with probability $p = [1 + e^{-\beta(f_B - f_A)}]^{-1}$, where β denotes the intensity of selection. An individual can either adopt the new belief, if he does not already, or he can stop believing it, if he comes across a neighbor who does not believe it. This will also be true for the next model, VM.

The simulation is stopped after 10^3 passages on all nodes, or when there are no more ignorants or spreaders of information. [9]

Voter Model (VM). The Voter Model is a simple model of opinion formation and dissemination. It works in the following way: each individual in the network randomly chooses one of its neighbors (a voter), and then proceeds to accept and adopt the neighbor's opinion with probability $p = 1$. [8]
As it happened in the PD model, the simulation is stopped after 10^3 passages on all nodes, or when there are no more ignorants or spreaders of information.

Susceptible-Infected-Recovered Model (SIR). Diffusion of ideas in a population has some aspects comparable with the propagation of infectious diseases. Thus, some models used to model disease spreading can also help us model information transmission. As mentioned above, each node can either be Susceptible/Ignorant, Infected/Spreader or Recovered/Stifler. A node can go through these three stages, in this order: If he's Susceptible, he can then become Infected (and actively spreading the disease/opinion), and afterwards become Recovered, where he's no longer susceptible nor spreading the disease. An idea begins in a single individual and spreads to its neighbor at a rate α. After some iterations as a Spreader, the individual gets tired of sharing the information (recovers) at a rate β. The simulation will stop when there are no more Spreaders active in the network. [1]

4.2 Network Configuration

The four network classes used all present similar low levels of clustering and increasing variance of degree distribution:

Homogeneous Small-World (HoSW). A Watts-Strogatz model was used to obtain Homogeneous Small-World. The probability of swapping the ends of pairs of randomly chosen links of the regular ring was set to $p = 0.1$, in order to keep the homogeneity.

Barabási and Albert (BA). The Scale-free Networks were created using a Barabási and Albert Model with preferential attachment.

Exponential Network (Exp). Exponential networks are obtained by using a similar model to the Barabási and Albert but random attachment is used instead of preferential attachment.

Heterogeneous Small-World (HeSW). Heterogeneous Small-World were built using the Watts-Strogatz model with $p = 1$ to create a irregular network in which all links are rewired.

4.3 Correlation Patterns

Denoting j as the number of individuals sharing the belief in a population of Z. The propensity $\delta_n(j/Z)$ measures the influence that one individual's belief has on the others at a distance n from him, relative to a random distribution of beliefs:

$$\delta_n(j/Z) = \frac{\epsilon_n(j/Z)}{\epsilon_n^{rand}(j/Z)} - 1 \tag{1}$$

$\epsilon_n(j/Z)$ is computed during each simulation. It results from averaging the probability that the neighbors at distance n of a node share the same belief as him. $\epsilon_n^{rand}(j/Z)$ is computed in the same way, with the same j/Z, however, the spreaders of information are randomly distributed.

The code to our implementation was done in Python using NetworkX.

It is to note that optimization of the algorithms was not a priority during the development.

5 Results

To evaluate the results, we used the normalized correlation values - δ_n/δ_1. Additionally, the critical network distance n_c represents the largest

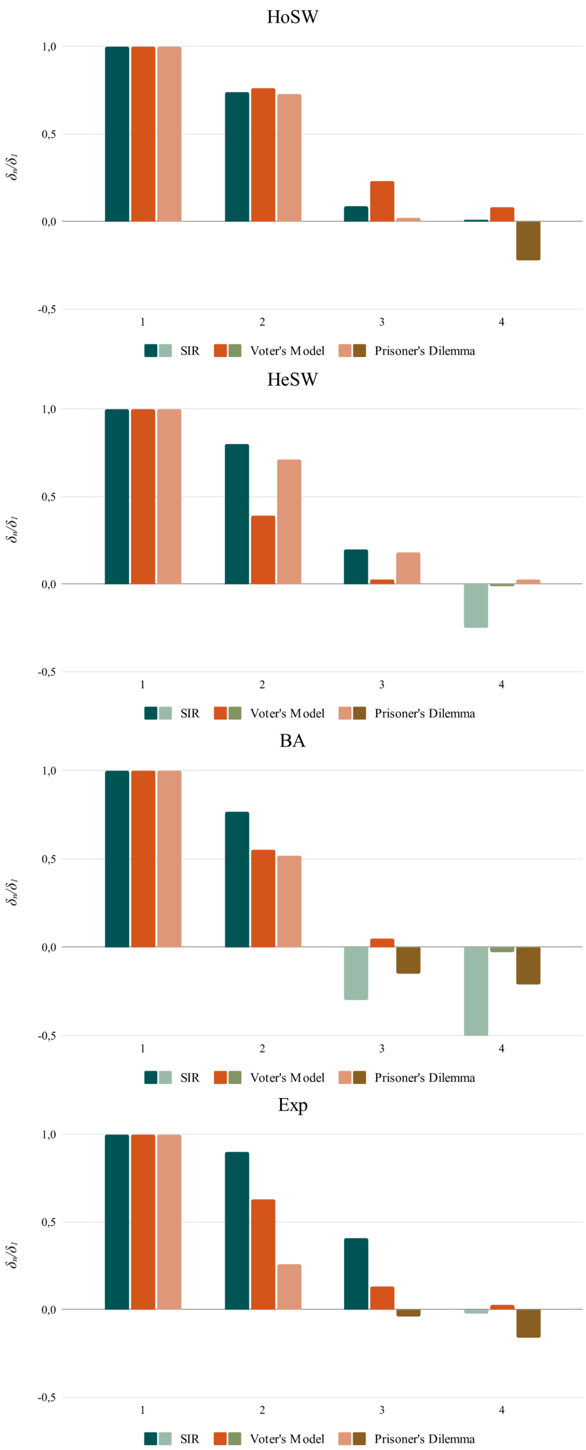

Figure 2: δ_n/δ_1 in relation to the increase of social distance *n* for simulation of SIR, Voter Model and Prisoner's Dilemma, taking place in Homogeneous Small-World (first graph); Heterogeneous Small-World (second graph); Barabási and Albert (third graph) and Exponential (fourth graph).

number n (network distance) for which we still have positive correlation values (δ_n/δ_1) ($n_c = 3$ represents a degree of influence equal to 3).

Figure 2 contains the results from running the different network configuration with $Z = 10^3$. For each graph of each network (HoSW, HeSW, BA and Exp) the results from the three models are shown (SIR, Voter Model and Prisoner's Dilemma). The propensity (δ_n/δ_1) reduces according to the increase of distance n, reaching a point where the correlation is not significant and, therefore, the value becomes negative (critical point n_c). The value n_c is not systematically equal to 3, as the value varies between 2 and 5, just as was shown in the *Origin of Peer Influence in Social Networks* ref. [7]. Some variables can affect the outcome, such as the average degree and the degree distribution. However, it's undeniable that all of the simulations follow the same kind of trend.

In case of a low average degree $\langle k \rangle$ and a high diameter, the networks became very sparse. The sparser the network, the more a node seems to increase the reach of his influence. This manifests in having a n_c that is greater than 3. Figure 3 shows a simulation of PD in a HeSW network with a very low average degree, $\langle k \rangle = 2$.

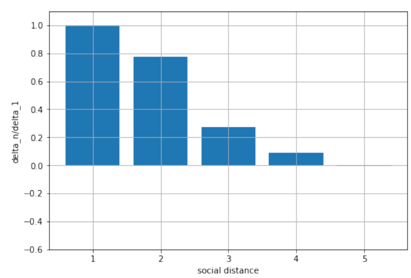

Figure 3: ($n_c >= 5$) δ_n/δ_1 in relation to the social distance n in a network with low average connectivity (Prisoner's Dilemma in a HeSW).

In the paper ref. [7], its shown that n_c stabilizes at that value 2 if we keep increasing the network connectivity, $\langle k \rangle$. Figure 4 represents a Homogeneous Small-World network with high degree, $\langle k \rangle = 10$, in a Voter Model simulation.

In the calculus of the neighbors, we only consider the shortest path, meaning that the distance n is the smallest distance between two nodes. For that reason, when we increased the network connectivity, by increasing $< k >$,the shortest paths get reduced, and so is the reach and influence of a node. We believe that the fact the value stabilizes at $n_c = 2$ might be caused by this.

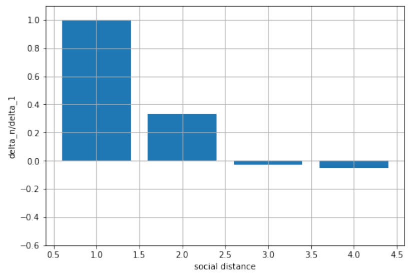

Figure 4: ($n_c = 2$) δ_n/δ_1 in relation to the social distance *n* in a network with high average connectivity, $\langle k \rangle = 10$ (Voter Model in a HoSW)

In conclusion, our results seem to be in agreement with the results of *Origin of Peer Influence in Social Networks* ref. [7]. Regardless of the type of social interaction or population organization, the same pattern of peer influence seems to always prevail. We modelled three different dynamical processes and they all showed similar degrees of the impact an individual makes on others around him.

Acknowledgments

We would like to thank professor Francisco Correia dos Santos, for both the development of the paper we recreated and for all we learned in the lectures and professor Francisco Miguel Enxerto Sena, for all the help, patience and guidance during the laboratory classes.

References

[1] Luís M.A. Bettencourt, Ariel Cintrón-Arias, David I. Kaiser, and Carlos Castillo-Chávez. The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications*, 364:513–536, 2006.

[2] Nicholas A. Christakis and James H. Fowler. The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, 357(4):370–379, 2007. PMID: 17652652.

[3] Nicholas A. Christakis and James H. Fowler. The collective dynamics of smoking in a large social network. *New England Journal of Medicine*, 358(21):2249–2258, 2008. PMID: 18499567.

[4] James Fowler and Nicholas Christakis. Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the framingham heart study. *BMJ (Clinical research ed.)*, 337:a2338, 02 2008.

[5] James Fowler and Nicholas Christakis. Connected: The surprising power of our social networks and how they shape our lives, little, brown, new york, ny. 353 pages. 2009.

[6] Zoya Gervis. Average american has this many actual friends, study determines. `https://www.foxnews.com/lifestyle/american-number-actual-friends-study-determines`, May 2019.

[7] Flávio L. Pinheiro, Marta D. Santos, Francisco C. Santos, and Jorge M. Pacheco. Origin of peer influence in social networks. *Phys. Rev. Lett.*, 112:098702, Mar 2014.

[8] V. Sood, Tibor Antal, and S. Redner. Voter models on heterogeneous networks. *Phys. Rev. E*, 77:041121, Apr 2008.

[9] György Szabó and Csaba Tőke. Evolutionary prisoner's dilemma game on a square lattice. *Phys. Rev. E*, 58:69–73, Jul 1998.

[10] Johan Ugander, Brian Karrer, Lars Backstrom, and Cameron Marlow. The anatomy of the facebook social graph. *arXiv preprint*, 1111.4503, 11 2011.