Developing a New Model to Characterize Cooperation in a Network

Yu Jin Lee, Xiao Zhou

1. Introduction

When and how do people cooperate? Cooperation is an important social phenomenon, whereby dynamic players interact with multiple agents in a complex network, playing mixed strategies. The theoretical front in evolutionary game theory has modeled cooperation as a strategy in a pairwise prisoner's dilemma played in a simple graph like rings and lattices [4, 5, 6]. In particular, cooperation only evolves as a dominant strategy under specific conditions on the average degree of a node, size of neighborhoods, and other environmental factors.

In this paper, we developed two network models to study cooperation in complex and realistic graphs, undirected and weighted character graphs based on the script of a popular TV show, Game of Thrones. Inspired by the notion of PageRank, we introduce a "generosity factor," a continuous variable ($g \in [0, 1]$) that represents the level of cooperation in a player's mixed strategy in a given graph. We iteratively simulated the convergence of generosity factor and wealth across various parameters and found that the rankings of characters with respect to both wealth and generosity are closely related to the ranking of characters based on their PageRank score. This suggests that the network structure has a significant impact on the characters' wealth and strategies.

2. Related Work

2.1 The Game of Game of Thrones in Reading contemporary serial television universes: A narrative ecosystem framework. 2018. Beveridge A, Chemers M

Summary

This article is an example of using network analysis to seek insights into a complex society in the TV series Game of Throne which is a narrative-style novel portrayed by plentiful characters and interweaving plotlines. The authors perform an analysis of the character's network to reveal the protagonists in each season using a character centrality diagram (CCD). The CCD contains three components: PageRank, Eigenvector Centrality, Weighted Degree. PageRank is a measure of the character's agency (the character's contribution to the plot and the ability to change the plot); Eigenvector Centrality measures the character's prestige (how powerful a person's network is; i.e. connectivities to other important characters); and Weighted

Degree represents the number of interactions with other characters. All three measures are normalized to a value between 0 and 1 which is easy to visualize on a 2-D plot with hue. Using CCD the authors successfully identified the protagonists in the analyzed seasons. Two interesting things are discovered. First, the three measures tend to be aligned in later seasons in which the characters are fully developed and the storylines became complete. Second, the rise and fall of the characters' positions over time can be seen in the CCD plots with one exception: the Kingslayer Jaime Lannister whose journeys seem to run through the whole series.

Critique

Although the results are interesting, they do not reveal new insights about the plotline. In particular, the TV series is known for having an ensemble of protagonists, so different viewers have different ideas on the most important character, and this is reflected by the lack of convergence of the three measures of CCD in most cases. Nonetheless, the authors demonstrate the richness of this dataset and the usefulness of network analysis to draw insights from graphs.

2.2 A simple rule for the evolution of cooperation on graphs and social networks. 2006. Ohtsuki H, Hauert C, Lieberman E, Nowak M., Nature

Summary

Cooperation is essential to human society. Although natural selection favors defectors over cooperators in unstructured populations, this paper reviews findings that in structured graphs, which resemble the actual biological systems more closely, natural selection favors cooperation. Cooperation is measured as a strategy in a pairwise prisoner’s dilemma game, whose nash equilibrium is a defection from both players. Simulations and mathematical calculations on structured graphs, such as cycles, spatial lattices, random regular graphs, and scale-free networks, suggest that the condition for the preference for cooperation is such that the ratio of the benefit of cooperation b to the cost of cooperation c higher than the average degree k. ($b/c > k$).

Critique

This paper introduces a framework to study cooperation, an essential social phenomenon, with mathematical rigor. However, since it is a theoretical contribution, the models are fairly simple and random. The authors suggest empirical and observational research to test the theory as further work. On a more theoretical side, the study of the “direct or indirect reciprocity” and the relationship between the network structure and the level of cooperation as further areas of research. We claim that our dataset will allow us to conduct both observational research to an extent and also the theoretical questions on the interaction between the structure and the level of cooperation. Furthermore, models in this paper assume homogeneity of population: all players are equivalent in terms of the capacity of being cooperator or defector only and the benefit and cost are parameters of the network, not individuals; however, in real life, humans tend to play mixed strategy with varying degrees of altruism and personalities. This limitation motivated an alternative definition of cooperation in our model that represents each character's propensity to give, rather than a fixed strategy with fixed cost and benefit for the entire network.

2.3 Network Games. 2010. A. Galeotti, S. Goal, M.O. Jackson, F. Vega-Redondo, and L. Yariv The Review of Economic Studies

Summary

Economists have modeled a wide range of social and economic interactions as non-cooperative games, in which each player plays different strategies to maximize self-interest according to a given payoff function. This paper introduces a framework to study games in a graph such that the payoff function depends on the actions of the player and his/her neighbors. The main challenge in this domain is that there are multiple equilibria even for the simplest games. The proposed framework fixes one Nash equilibrium with the incomplete information assumption: each player only knows his/her degree, but not the identity of their neighbors and their degrees and the strategies are determined before the network is "realized," such that this incomplete information serves as a constraint on determining one's strategy. The two types of games studied are strategic substitutes and strategic complements. The former kind is a game in which a player's likelihood of choosing an action decreases with the increased adoption of the action by his/her neighbors, and the latter is the opposite. Under this framework, several propositions about the existence of an equilibrium in a network satisfying specific payoff conditions and modified networks.

Critique

Although it is important to introduce constraints to avoid the multiplicity of equilibria, the incomplete information assumption is not generalizable to many situations in reality. It is more plausible that a player personally knows his or her neighbors and has beliefs about their degree, which then informs the player's action. For instance, an agent may favor neighbors with high degree centrality compared to the less connected counterparts in anticipation of favors from their wide network. Furthermore, while this assumption is crucial to fixing an equilibria, the a priori determination of strategies to the formation of networks is not suitable for studying repeated games or dynamic networks that evolve over time. To complement this weakness, we simulate a generalized form of trust game in which a portion of the player's wealth is invested to and returned from its neighbors on a realistic social network.

3. Data

Our data set is a character network generated from the fan-generated scripts of the TV show Game of Thrones found in genius.com [1]. The weight of an undirected edge between two characters is the sum of the number of the following types of interactions between the two:

appearance in the same scene, appearance in the same stage direction, exchange of dialogue, one character mentioned by the other, and two characters are talked about together by a third character. The authors note that there is an emphasis on verbal interaction, represented by dialogue, over physical interaction, represented by stage direction, which renders extended action scenes less significant than dialogue. Nevertheless, given the importance of verbal interaction in the evolution of storyline in animated narratives, this is a reasonable decision. On average, there are ~120 characters and edges on the order of a few hundreds for a network of one season. The weight of edges shows a fat-tailed distribution, with less than 10 edges with weights higher than 100 while most of the edges have a weight below 20.

In addition to the network, we manually curated a list of characters' death in each season based on the death data published online. The number of deaths in a season increases as the season progresses, with 59 deaths in season 1 to more than 4000 deaths in season 8; however, trivial characters who do not have significant interaction with other characters are not included in our character graphs.

4. Models and Metrics

4.1 Generosity Model 1

In the first model, we assume that a character i has an initial wealth $w_i^{(0)}$, representing the fitness score in an evolutionary game theory framework, and a generosity factor $g_i^{(0)}$. At every iteration, each character distributes a $g_i^{(k)}$ fraction of wealth to his/her neighbors for trading and receives wealth from his or her neighbors of $B(w)$. $B(\cdot)$ is a beneficiary function, representing the benefit-received-to-wealth conversion. For simplicity, we will set $B(w) = \alpha \cdot w$, where α is a wealth converting constant (> 0). If $\alpha > 1$, more wealth is generated from the wealth received, suggesting a non-zero-sum game setting, while $\alpha < 1$ represents a setting where wealth diminishes as it is distributed. After each iteration, the total wealth of a character is the sum of the remaining wealth after distribution and the converted wealth received from his or her neighbors. In summary, for all the characters, the wealth vector $w^{(k)}$ after iteration k is:

$$w^{(k)} = (I - D^{(k)})w^{(k-1)} + \alpha M D^{(k)} w^{(k-1)} = (I - D^{(k)} + \alpha M D^{(k)})w^{(k-1)}$$

Where

- I is the identity matrix
- $D^{(k)}$ is a diagonal matrix whose value at ii is the generosity factor $g_i^{(k)}$ for the person i
- M is the matrix similar to the PageRank matrix with $M_{ij} = d_{ji}/\Sigma_{k=1}^{n} d_{jk}$, d_{jk} is the element in the adjacency matrix of a Game of Throne graph
- α is the wealth converting constant

We simulate the process by updating $g^{(k)}$:

$$g^{(k)} = g^{(k-1)} + \delta\{[w^{(k)} - w^{(k-1)}]/w^{(k-1)}\}$$

i.e. $g^{(k)}$ is adjusted by each person's return rate $[w^{(k)} - w^{(k-1)}]/w^{(k-1)}$ at each iteration. δ is a small number and after experimenting it is set at 0.05. To avoid "spider trap" and "sink loop", $g^{(k)}$ will never go below 1e-12.

We realized that in this model updating generosities is very similar to adjusting the edge-weight of self-loops and that in this model M does not change during the iterative process, which means that each person adjusts the total portion of wealth distribution, and the individual feedback from previous iterations are not incorporated. We then developed model 2 to allow a feedback loop for every individual pair.

4.2 Generosity Model 2

In model 2, we wait for w to converge and then let each character adjust their investment rate for each of their neighbors based on the return ratio. We repeat this for a large number of iterations, or until g and w converge.

Let J be the matrix whose elements are all the 1's.
Let I be the identity matrix.
Let $M_{ij} = max(d_{ji},\ 1e-12)/\Sigma_{k=1}^{n} max(d_{jk},\ 1e-12)$, d_{jk} is the element in the $g^{(0)}$-adjusted adjacency matrix of the Game of Throne graph

For iteration $k = 0...K$:

- Let $D_w^{(k)}$ be a diagonal matrix whose element ii is $w_i^{(k)}$.
- Let $C^{(k)} = M^{(k)}D_w^{(k)}$, $C_{ij}^{(k)}$ is the cost of wealth from node j to node i.
- Let $T^{(k)} = (\alpha J - \alpha I + I)M^{(k)}D_w^{(k)}$, $T_{ij}^{(k)}$ is the wealth node i received from node j.
- Let $E^{(k)} = E^{(k-1)}/2\ +\ T^{(k)}/2$, which is the time-damped average of $T^{(k)}$. We use a time-damped average instead of $T^{(k)}$ to make the calculation more stable because $T^{(k)}$ may swing back and forth at neighboring iteration. In addition, this is also a simple emulation of a real-life investment strategy.
- Let $R^{(k)} = (E^{(k)\prime} - C^{(k)})/C^{(k)}$, $R_{ij}^{(k)}$ is the return ratio of the wealth invested by node j to node i. $E^{(k)\prime}$ is the transpose of $E^{(k)}$. (After wealth converges)
- Wealth vector $w^{(k+1)} = \sum_j T_{\cdot j}^{(k)} / \sum_{ij} T_{ij}^{(k)}$, i.e. row-wise summation of $T^{(k)}$ and normalized to a stochastic vector.
- Generosities vector $g^{(k+1)} = \sum_i [M^{(k)} \circ (J-I)]_{i\cdot}$, i.e. column-wise summation of $M^{(k)} \circ (J-I)$, $\circ$ is element-wise multiplication.

- If $w^{(k)} - w^{(k-1)} < 1e-12$ (wealth vector converged), (1) update $M^{(k+1)} = (\delta R^{(k)} + J) \circ M^{(k)}$, element-wise replace <=0 to 1e-12 in $M^{(k+1)}$, and (3) then column-wise normalize $M^{(k+1)}$ to a stochastic matrix, else $M^{(k+1)} = M^{(k)}$.

For simplicity, we fixed $g^{(0)}$ at 0.5 by assuming all the characters are equally generous/selfish in the beginning. Because we let $w^{(1)}$ converge first, initialization of $w^{(0)}$ is less relevant in model 2. Now the parameters of the system include α and δ. We found that when δ was small (~1) $g^{(k)}$ barely changes. After experimenting, we set $\delta = 1.3$. We analyzed a range of α settings on the impact of society. **All our analyses are done using Model 2.**

4.3 Metrics Used

Average Precision @K (AP@K)

AP@K allows us to compare the final rankings with respect to wealth and the generosity factor and the ranking based on the PageRank of the initial g_0. It is defined as the following:

$$AP@K = \sum_{k=1}^{n} \frac{\{test\ ranking\ top\ k\ items\} \cap \{truth\ ranking\ top\ k\ items\}}{nk}$$

n is the number of nodes in the network.

The Gini Coefficient

It is used to measure economic inequality for the final wealth vector w. The Lorenz curve is the wealth distribution curve obtained by plotting the population percentile by income on the horizontal axis and cumulative income on the vertical axis. The perfect equality line is the 45-degree diagonal line. The Gini Coefficient of 0 indicates perfect equality, and 1 a maximal inequality.

The Gini Coefficient = $\frac{\text{the area below the line of perfect equality} - \text{the area below the Lorenz curve}}{\text{the area below the line of perfect equality}}$

Graph Transitivity

Transitivity $T = 3\frac{\#\ of\ triangles}{\#\ of\ triads}$, a "triad" is a structure of two edges with a shared vertex.

Graph Average Clustering Coefficient

Average Clustering Coefficient $C = \frac{1}{n}\sum_{i}^{n} C_i$, where $C_i = \frac{2e_i}{k_i(k_i-1)}$, e_i is the number of edges between the neighbors of node i.

Graph Connectivity

Graph connectivity is the size of the largest connected component.

Spearman's Rank Correlation Coefficient

$r_s = \frac{cov(rg_x, rg_y)}{\sigma_{rgx}\sigma_{rgy}}$, where $cov(rg_x, rg_y)$ is the covariance of the rank variables rg_x and rg_y, σ_{rgx} and σ_{rgy} are the standard deviations of the rank variables.

5. Results and Discussion

5.1 General Observation

We first performed 100,000 iterations using the Generosity Model 2 and discovered that (Figure 1):

- If $\alpha \geq 1$, g again converges to all 1s and w converges to a vector related to the initial PageRank. The Gini Coefficient is stable.
- Interestingly, if $c.a.\ 0.95 < \alpha < 1$, g still converges at all 1s and the Gini Coefficient is the lowest.
- If $\alpha < 0.9$, the wealth is concentrated at the people who have the highest initial PageRank scores. Although the values of the converged g is not the same for each α, the ranking is relatively converged and people who have less wealth are more generous, as shown by the AP@K of ascending g to descending w in Figure 1a.

a b

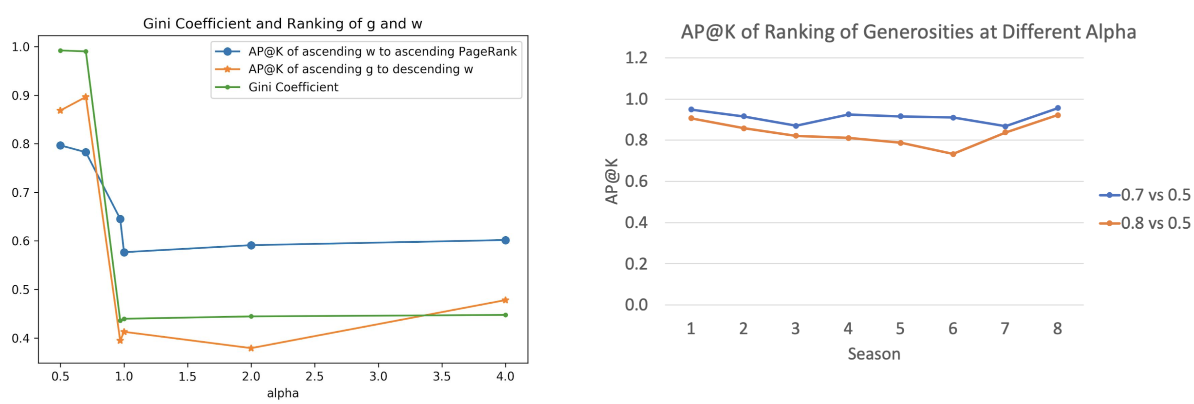


Figure 1. (a) Gini Coefficient and ranking of converged generosities and wealth compared to the initial ranking of PageRank (100000 iterations). (b) AP@K of the ranking of generosities at $\alpha = 0.7$ vs $\alpha = 0.5$ (blue line) and $\alpha = 0.8$ vs $\alpha = 0.5$ (orange line) for the 8 seasons graphs.

To illustrate this finding, we choose three characters NED, HOUND, and HIGH_SEPTON whose has high, medium, and low initial PageRank scores, respectively, to showcase the changes in w and g through the iterations with varying parameter (Figure 2). It is clear that when $\alpha = 0.5$, representing a partially zero-sum game, the wealth is transferred from

HIGH_SEPTON and HOUND, the lower and middle characters, to NED, the already rich character. HIGH_SEPTON, the lower class character becomes maximally generous because his wealth is so little that any investment to the rich, who despite being minimally generous still has significantly greater wealth and thus returns an amount that significant to the lower class, is desirable. This result is expected given our design choice to force a minimum requirement of g = 1e-12 for all characters, and we can reasonably deduce that without the requirement, the generosity factor of the “rich”, who has no incentives to make an investment that has negative returns, will converge to 0. Interestingly enough, the observation that people become “less generous” as they become richer and thus more powerful is consistent with the psychology literature. When $\alpha = 1$, representing a win-win game, the wealth and generosity of all characters quickly stabilized, with everyone becoming maximally generous. Given the maximal generosity, the wealth converged to the values comparable to the initial wealth.

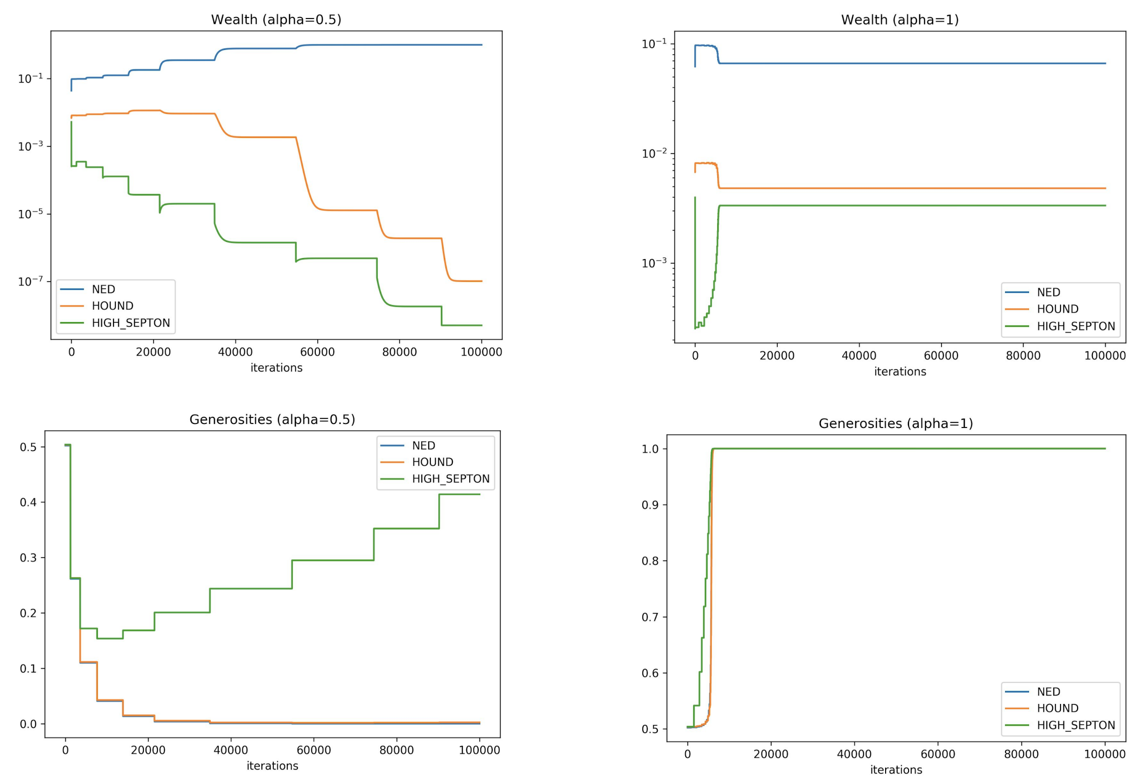


Figure 2. Changes of wealth and generosities over iterations when $\alpha = 0.5$ and $\alpha = 1$. The wealth is in the log-scale.

5.2 Effect of Number of Iterations

We then performed different numbers of iterations to reflect a more realistic condition that unusual events that allow people to adjust strategies and the wealth to be redistributed, such as wars and recessions, are temporally limited. Figure 3 shows that the number of iterations has strong effects when $\alpha > 2$. Shorter iteration time benefits the people with higher initial PageRank scores (final rank of wealth positive correlated with the initial PageRanks), which creates higher economic inequalities.

In addition, Model 2 allows the initial graph to evolve based on each character's "investment returns". We would like to compare the difference between the initial and the final graphs.

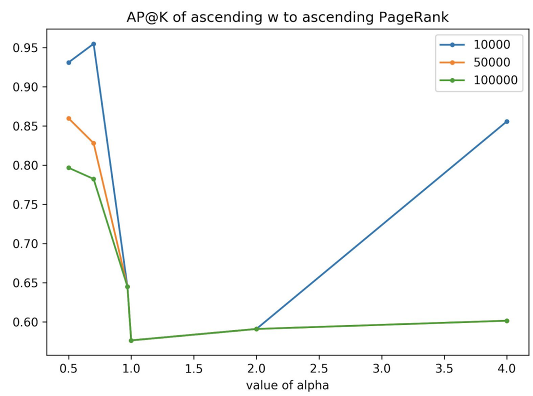


Figure 3. Effects of the number of iterations on the ranking of final wealth vs the ranking of the PageRank of the initial graph

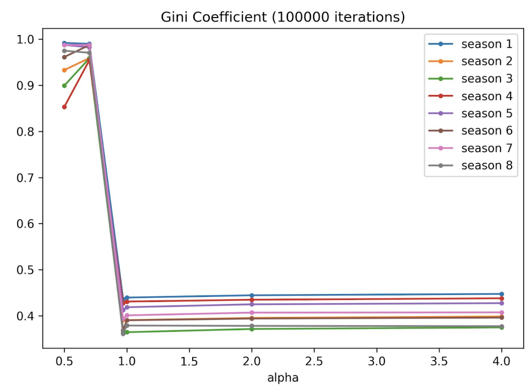
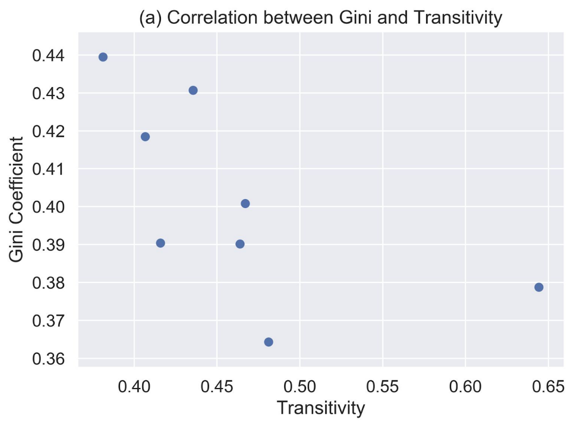


Figure 4. Gini Coefficient vs different α *values for the 8 seasons' graph. Notice the smallest value is at ~* $\alpha = 1$.

5.3 The Graph Structure Governing the Gini Coefficient

We discovered that for a given graph the lowest Gini coefficient after 100,000 iterations is reached when $\alpha = 1$; however, the minimum values are different across the graphs, which suggests that the final Gini coefficient is influenced by the graph's structural properties as well as the parameter we chose (Figure 4). Thus, we analyzed three graph properties: (a) transitivity (b) average clustering coefficient, and (c) connectivity. Although all three properties relate to the connectivity of the graph, high transitivity indicates that that the whole graph is well connected, while a high average clustering coefficient may result from a collection of well intra-connected clusters. During all calculations, the edge weights are ignored, and we fix the parameter $\alpha = 1$ to control for the effect of the parameter on the wealth gap.

As shown in Figure 5, there is a strong negative correlation between the Gini coefficient with transitivity (Spearman's R = -0.81; p-value = 0.015), but not with the other two measures of connectivity. This suggests that an overall well-connected network, as opposed to a network of tight clusters, is a foundation for a low wealth gap. In practice, this suggests that, under the condition of fair trading ($\alpha = 1$) of our Model 2, globalization, which increases connectivity of the network, benefits all nations and lowers the Gini Coefficient, while protected trading benefits the rich and well-connected countries only, increasing wealth gaps.



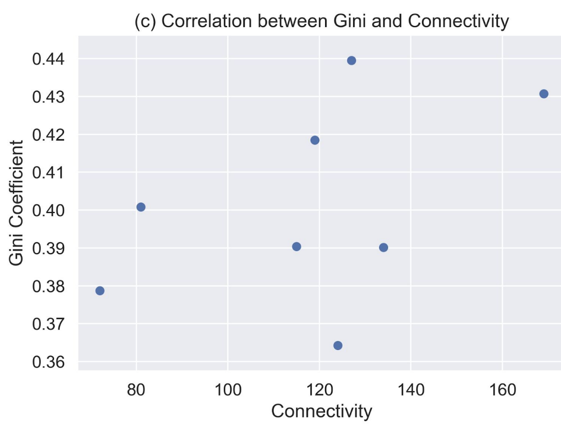
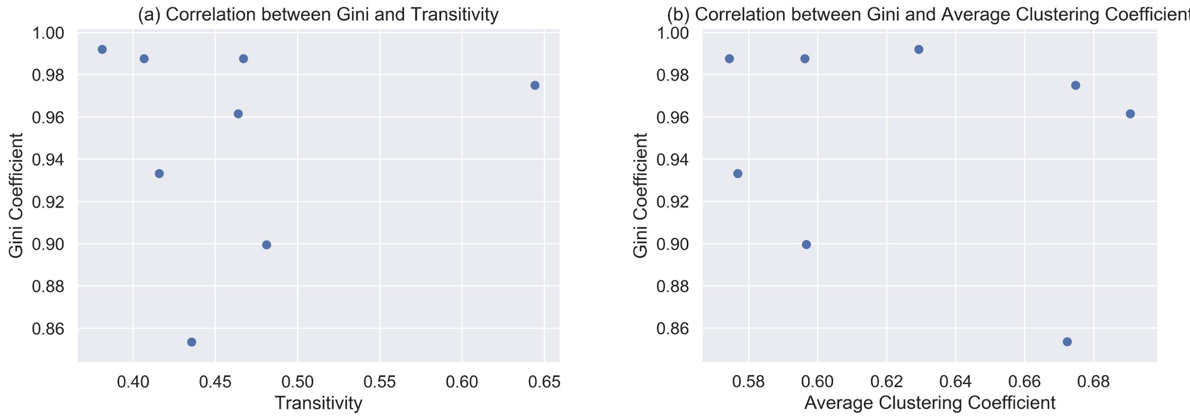


Figure 5. (a) Correlation between the Gini Coefficient and Transitivity. (The outlier at right is season 8). The Spearman Correlation Coefficient is -0.81 and the p-value is 0.015. (b) Correlation between the Gini Coefficient and the Average Clustering Coefficient. The Spearman Correlation Coefficient is -0.21 and the p-value is 0.61. (c) Correlation between the Gini Coefficient and the Connectivity The Spearman Coefficient=0.38 and the p-value is 0.35). $\alpha = 1$ *and* edge weights are not considered.

When $\alpha = 0.5$, which represents a more bleak world in which only half of the favors are returned, we did not find any statistically significant relationship between the Gini Coefficient and the measures of connectivity (Figure 6). This suggests that the parameter alpha dominates the network properties in determining the Gini Coefficient, which is expected. The parameter alpha in real life is likely to be between 0.5 and 1, so we cannot rule out the possible role of the network connectivity on determining the wealth inequality.



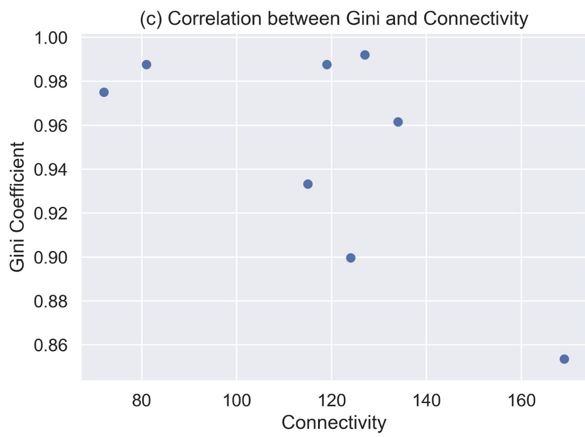


Figure 6. (a) Correlation between the Gini Coefficient and Transitivity. The Spearman Correlation Coefficient is -0.40 and the p-value is 0.32. (b) Correlation between the Gini Coefficient and the Average Clustering Coefficient. The Spearman Correlation Coefficient is -0.23 and the p-value is 0.57. (c) Correlation between the Gini Coefficient and the Connectivity The Spearman Coefficient=0.33 and the p-value is 0.41).

6. Conclusion and Future Work

In this paper, we've proposed a new model to study cooperation in a network inspired by the idea of PageRank. By simulating a modified version of public goods game, in which a chosen portion of a player's wealth is "invested" to its neighbors then returned according to a parameter alpha, we investigated the relationship between a player's position in the network, measured by the Pagerank ranking, its final wealth, and generosity. Furthermore, we studied various factors that influence the Gini Coefficient, which measures the level of wealth inequality.

Although alpha, the parameter of the game, is the dominant factor in determining the final Gini coefficient, when $\alpha = 1$, suggesting an honest investment, we found a statistically significant correlation between the Gini coefficient and transitivity, which is a measure of connectivity that considers connectedness in a global sense, but not with other measures of connectivity that are more local. Simulating games on a realistic network is a novel approach that has yielded interesting results and potential policy implications.

In our model 2, each node uses a relatively simple investment strategy by monitoring the investment return ratio and use it to allocate future investment. However, a more complex strategy can be implemented in the future. For example, each node can be a reinforcement learning model that is trained to adjust its investment based on the returns to maximize its wealth in the long run. We hypothesize that under such condition, a trading coalition may form depending on different values of α.

References

1. Beveridge A & Chemers M (2018) The Game of Game of Thrones in Reading contemporary serial television universes: A narrative ecosystem framework.
2. Galeotti A., Goyal. S, Jackson M.O., Vega-Redondo, F., Yariv L., (2010) Network Games. The Review of Economic Studies **77**, 218–244
3. Girvan M & Newman MEJ (2002) Community structure in social and biological networks. *Proc Natl Acad Sci USA* **99**, 7821–7826.
4. Hauert C & Doebeli M (2004) Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* **428**, 643–646.
5. Ifti M, Killingback T & Doebeli M (2004) Effects of neighbourhood size and connectivity on the spatial Continuous Prisoner's Dilemma. *J. Theor. Biol.* **231**, 97–106.
6. Ohtsuki H, Hauert C, Lieberman E & Nowak MA (2006) A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505.