# Course Project
# Complex System and Network Science

Federico Rullo

[1] University of Bologna
[2] (ID: 0001026401)
[3] federico.rullo@studio.unibo.it
Master's Degree in Artificial Intelligence

**Abstract.** Experiment-based work where we develop a simulation-based on the research done by Scott L. Feld, which supposed that individuals use the number of friends that their friends have as one basis for comparison- son to measure their adequacy in their social circle. The idea is to create a simulation of a friendship network using different models and observe and analyse how the network evolves based on different factors affecting the different individuals.

**Keywords:** Friendship Networks · Random Graph · Small World · Preferential Attachment · Game Theory

## 1 Introduction

Friendship is a fundamental aspect of human social interaction and can provide a sense of satisfaction, belonging and support. However, friendships can also be used as a way for individuals to evaluate themselves and others, leading to feelings of inadequacy if the comparison is not favourable. To better understand this phenomenon, we are going to create a simulation based on the article published by Scott L.[1], which describes a real-world study on friendships.
In this study, data were drawn from Coleman's study [2] *The Adolescent Society*, where he collected the number of friendships among students in 12 high schools.

To better understand this phenomenon we analyze an example extracted from Coleman's work, Figure 1 shows friendship with the number of friends of each individual and, in parenthesis, the mean of friends of each friend. Only *Sue, Alice* and *Carol* have at least the same number of friends as the number they will compare to or more, while every other individual will feel inadequate.
We observe the distribution of the number of friends and the distribution of friends of friends. The latter, however, takes into account some individuals more than one time each individual's friend contributes to the final friends' average as many times as she has friends.
Whenever each individual compares their number of friends with the average number of friends' friends the comparison is unfair: two different distributions are taken into account.

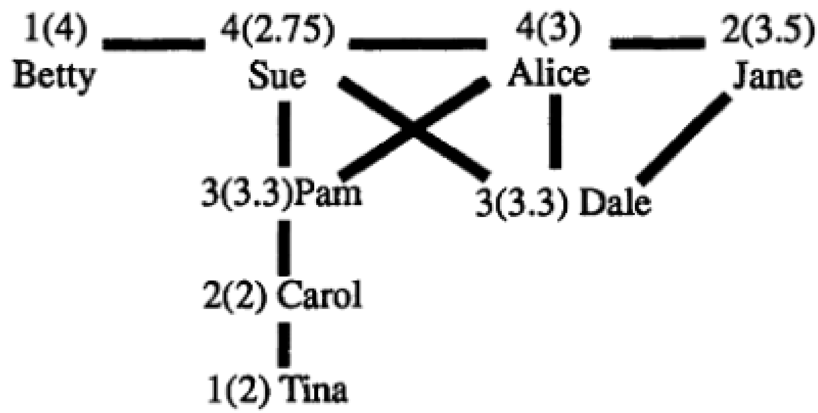


**Fig. 1.** Friendship examples based on data from Marketville High School

There is a simple relationship between the distributions. The average number of friends of each individual is simply

$$AVG_f = \frac{\sum_{i \in I} x_i}{N} \tag{1}$$

Where $x_i$ is the number of friends of individual $i$ and $N$ is the total number of individuals in the network, and $I$ the set of individuals.
On the other hand, the average number of friends' friends of an individual is

$$AVG_f f = \frac{\sum_{i \in I} x_i^2}{\sum_{i \in I} x_i} \tag{2}$$

where $x_i^2$ is the squared sum of friends' friends of individual $i$ and $x_i$ the total number of friends of $i$.
The simulation aims to observe how the pool of friends of an individual will change over time, and how different initial configurations and models impact the formation of new friendships. In particular, we consider four models: Erdos-Renyi, Watts-Strogatz, and Barabasi. Finally, a model based on game theory, where individuals may decide to be friendly or not, has been implemented using the initial configuration of Erdos and Watts to see how it affects the formation of new friendships.

## 2 Proposed Models

To make the simulation more dynamic and realistic, we will assign each individual in the network a popularity parameter that reflects their attractiveness as a potential friend. This allows for variation within the simulation and allows for the observation of different behaviours within the same network.

The Erdos-Renyi model is a random network model that assumes that each individual has an equal probability of forming a connection with any other individual in the network. This model is useful for understanding the formation of friendships in large, diverse populations where there are no clear preferences or biases. The Watts-Strogatz model is a small-world model that combines the characteristics of both regular and random networks. It has a high level of clustering, like a regular network, but also a high level of randomness, like a random network. This model is useful for understanding the formation of friendships in small, homogeneous communities where there are strong ties among neighbours, but also some level of exploration and diversification. The Barabasi model is a scale-free network model in which a few individuals have many connections and the majority have few connections. This model is based on the idea of preferential attachment, in which individuals are more likely to form connections with those who already have many connections. This model is useful for understanding the formation of friendships in hierarchical or oligarchical societies where social status or power plays a significant role.
Finally, the game theory model will start from a given configuration between Erdos and Watts and use a combination of simple rules to determine the formation of new friendships.
This model is useful for understanding the formation of friendships in societies where there are explicit or implicit rules or expectations governing social interactions.

### 2.1 Model Implementation

To implement the simulation, we will use the Netlogo platform, version 6.3.0, which is a widely used tool for modelling and simulating complex systems. The different variables used, which are shared among the models are the total number of individuals (N), which can vary from 10 to 100 individuals, to accommodate some of the models that require a high number of individuals and to have a better view of the behaviour of each model when a high number of individuals are present in the network, and a list of probabilities to have better control over the simulations, and test different configurations, which are shared by all models. These probabilities are:

- The probability of an individual making new friends
- The probability of an individual losing a friend
- The probability that an individual will be popular

These probabilities will dictate the behaviour of the individuals during the simulation, and how connections will evolve.
An additional variable is a selector, which allows us to run the Barabasi model using popularity or not, this will allow us to see how much difference popularity makes on friendships in preferential attachment models.
All these parameters are accessible in the setup phase of the model in form of sliders in the Netlogo's interface.

To observe the simulation we use three different components: the world window, which allows us to observe the network's evolution in real time, and two histogram plots, which give us insight into the average number of friends and the average number of friends of friends of each individual.

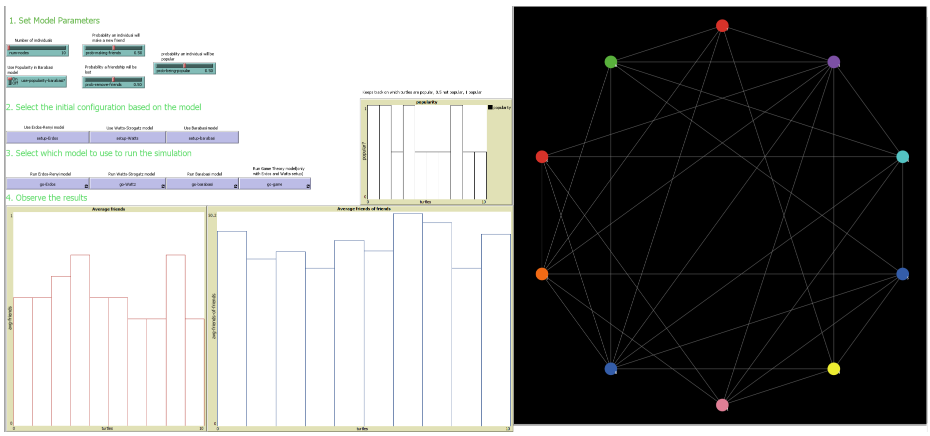


**Fig. 2.** Netlogo Starting interface

## 3 Model Evaluation

The purpose of this project is to provide a simulation and evaluate and compare the models described for simulating friendship networks, to see how different initial configurations will modify the network behaviour and which of the presented models is optimal to run such a simulation. The model evaluation will be done by observing the network in real-time, to see if it becomes too chaotic, and by plotting the histogram of the average number of friends and friends of friends for each individual in the network. The values for these histograms will be updated as the model runs, allowing for the observation of how the model changes over time. Here we give a brief explanation of the models and how they are configured for this simulation.

### 3.1 Erdos-Renyi

The Erdős–Rényi model is a statistical model for random graphs that generates a graph by starting with a set of n vertices (individuals) and adding edges between them randomly with probability p. This model is often used to study the properties of large networks, such as social networks, where the probability of two people being friends might depend on external factors like shared interests or geographic location. One of the key properties of the Erdős–Rényi model is

that as the number of vertices increases, the probability of the graph containing a giant connected component (a large group of vertices all connected) approaches a certain value. This value depends on probability p and can be calculated using the formula: $P(giantconnectedcomponent) = 1 - (1 - p)^n$.
For this simulation, we are going to use this model as an initial configuration of the network, but also to run the simulation, this way we can see how the model behaves over time and how the model's initial configurations affect other models, like the game theory model. The Erdős–Rényi model is modified to include an initial probability that an individual will be popular, drawing higher connections and when running, a probability that an edge between two individuals will be deleted, which introduces noise to the model and allows for the observation of its stability, this probability is affected also by the popularity of the individuals.

### 3.2 Watts-Strogatz

The Watts–Strogatz model is a random graph model that generates graphs with small-world properties. It starts with a ring lattice (a graph where each vertex is connected to its nearest neighbours) and rewires the edges with some probability p, resulting in a graph with a mix of regular and random connections. This model is often used to study social networks, where regular connections represent close friendships or familial relationships, and random connections represent more distant or casual relationships. One of the key properties of the Watts–Strogatz model is that it generates graphs with high clustering and small average path length, which makes it a good fit for representing small-world networks.
The Watts–Strogatz model has been modified for this simulation, to include the popularity of nodes, popular nodes have a higher rewiring probability than non-popular ones.

### 3.3 Barabàsi-Albert

The Barabàsi-Albert model is a random graph model that generates scale-free networks. A graph is generated by starting with a small number of vertices (m0) and adding new vertices one at a time. Each vertex is connected to an m existing number of vertices chosen uniformly at random with replacement. This process is repeated until the desired number of vertices is reached.
The BA model is often used to study the structure of networks, such as social networks, where the vertices represent individuals and the edges represent relationships. In this context, the BA model can be used to simulate the growth of a friendship network over time.
The Barabàsi-Albert model uses a mechanism called Preferential Attachment by which nodes in a network tend to connect to other nodes that are more popular or have more connections. This process can lead to the emergence of scale-free networks, which are characterized by the fact that a small number of nodes have a large number of connections, while the majority of nodes have only a few connections.
The model, like the previous ones, includes a popularity parameter that decides

if a newborn individual is popular or not, and the weight of the popularity parameter on new connections.
Finally, a Boolean parameter has been integrated to deactivate the added popularity, this way we can see how much the popularity of an individual will influence their connections.

### 3.4 Game Theory

The Game Theory model is a useful tool for understanding how individuals make decisions about forming friendships based on the expected benefits and costs of such relationships. In the context of a friendship network, the model considers the "friendliness" of an individual, which represents their willingness to form new friendships. This parameter is influenced by the relative importance that an individual places on the benefits and costs of friendship. For example, an individual who values the social support and companionship provided by friendships may have a high friendliness parameter, while an individual who is more concerned with the time and effort required to maintain friendships may have a lower friendliness parameter.
The model can be used to simulate the formation of friendships over time by iteratively updating the friendliness parameters of individuals based on the outcomes of their interactions with others. For example, an individual who consistently experiences positive outcomes from their friendships may have their friendliness parameter increase, while an individual who consistently experiences negative outcomes may have their friendliness parameter decrease. The model can be evaluated by comparing the number of friendships formed under different assumptions about the benefits and costs of friendship, and by examining how the model changes over time.
In addition to the friendliness parameter, the model will also consider a popularity parameter, which represents an individual's popularity or social status within the network. This parameter may influence the likelihood that an individual will form new friendships, with more popular individuals being less likely to isolate themselves than non-popular ones.
To introduce noise and stress into the model, some friendships may be deleted from the network given a probability. This allows for the observation of how the model performs under different conditions and can help to identify any weaknesses or limitations in the model.

## 4 Experiments

Three different configurations for the models will be carried out, to test how they will affect the initial state of the network for each of the models and how, during runtime, they will affect the simulation outcome.
We also increment the number of individuals with each configuration.

### 4.1 Configuration 1

*num-of-nodes: 10, prob-making-friends: 0.5, prob-removing-friends: 0.5, prob-being-popular: 0.25*
In this configuration, we start with a small batch of individuals, we aim to study how the models will behave if the probability of forming new friendships and losing existing ones were equal, while also limiting the probability that a popular individual is born to around 25%.

**Erdos-Renyi** In this model we expect popular individuals to start with more connections than non-popular ones, we also expect that, by running the model, connections between individuals will form and be deleted at random with popular individuals having the most number of connections between nodes, and maintaining their connections, in case of deletion, making it hard to reach a fully connected network, but having giant components in the form of popular individuals.
However, as the simulation progresses, the network becomes fully connected after approximately 100 ticks. This results in the average number of friends and friends of friends being equal for all individuals in the network regardless of popularity, providing no insight into the evolution of friendship networks over time. This can also be caused by the small number of individuals in the network.

**Watts-Strogatz** The initial configuration of the model is predefined for all configurations, individuals will be connected to their nearest neighbours at a distance of 2.
During the run time, we expect the popularity factor to influence how often an individual will swap their connections, hence having a large number of weak connections, while non-popular individuals have a low rate of change and maintain strong connections with similar individuals.
By using this model to run the network we can observe that this is the case, this can be also observed in their average number of friends which is higher than popular individuals, who in return will have a higher average in the number of friends of friends.

**Preferential Attachment** The initial configuration of the model is made of only two initial individuals, where either of them can be popular or not. During run-time, we expect popular individuals to draw more connections, which means that we should observe giant components forming in the networks, while non-popular individuals should have a small number of friends, around 1 or 2 maximum.
If the popularity parameter is not used, then we expect new individuals to connect randomly to already existing ones. However, more interesting results are expected when the popularity parameter is utilized.
By running the model we can observe that new individuals are drawn by popularity, allowing us to see how popularity impacts friendship networks, hence running the models meets our expectations.

### 4.2 Configuration 2

*num-of-nodes: 40, prob-making-friends: 0.7, prob-removing-friends: 0.3, prob-being-popular: 0.15*
This time we configure the network to have a higher probability of making friends and a low probability of losing them, we also lower the impact of popularity on the network by making the creation of popular individuals rare, by setting it around 15%.
Finally, we increase the number of individuals in the network to observe its impact on the runtime.

**Erdos-Renyi** This time we expect the initial configuration to be completely randomized, with a low number of popular individuals, also since the probability that two individuals will be connected is rather high, we expect the network to start with a high number of giant components. During run time, we expect popularity to play a small part in individual connections, and, since the probability of losing friends is much lower than in the previous configuration we expect the network to reach full connectivity in a small set of ticks.
As expected, the network quickly becomes fully connected, with popularity having little to no impact on the connections between individuals. The resulting simulation is too chaotic to be informative and does not provide any insights into the behaviour of individuals in terms of changes in their friendship group.

**Watts-Strogatz** The setup difference with the previous configuration will be the low number of popular individuals. This time we expect individuals to form and maintain strong connections, while only a few popular individuals will change frequently their group of friends. After running the model, we observe that non-popular individuals do indeed form and maintain strong connections with similar individuals, we can also observe that their averages tend to stay within a certain range with little or no variation.
On the other hand, popular individuals have a higher ratio of change in their connections, this ratio is prominent in their averages, where they can span from a high number of friends to only a few individual connections.

**Preferential Attachment** With preferential attachment, we expect to observe a more random network compared to the previous one, but also a few popular individuals that will monopolize the future connections, hence we expect to still see a big impact from popularity even if the ratio of popular individuals compared to the number of nodes is rather low.
As expected running the model we can observe that connections are all drawn to a small group of individuals, which are the popular ones. On the other hand, we see that the network is less randomized than expected, with non-popular individuals having no more than one or two friends, and one of them is always a popular one.

### 4.3 Configuration 3

*num-of-nodes: 60, prob-making-friends: 0.3, prob-removing-friends: 0.7, prob-being-popular: 0.6*
Again we increment the number of individuals in the network to 60, a maximum of 100 can be reached but that can make the network hard to read, as the number of individuals in the network increases, we can expect to see the network become more complex and dynamic, with a larger number of connections being formed and broken over time.
With this configuration, we try to do the opposite, this time we expect to see a network where the probability of losing friends is much higher than the probability of making new friends, and where also the majority of the population is made out of popular individuals. This should result in a network that is constantly changing and evolving, with popular battling for connections and little or no connections between non-popular ones.
We can also expect to see a larger gap between the number of average friends and average friends of friends for popular individuals, as they will have a higher initial weight during setup and will be more likely to attract new connections during run time. As the number of individuals in the network increases, we can expect to see the network become more complex and dynamic, with a larger number of connections being formed and broken over time.

**Erdos-Renyi** Using this configuration we expect the network to start with a high number of giant components, based on the high number of popular individuals in the network, during run-time we expect the network to never settle on a fully connected network as opposed to the other configurations, with giant connected components continuously changing their connections while also maintaining their size, more similar to what we observed in the Watts-Strogatz model.
Running the model we can observe that this is the case, with a huge cluster of connected individuals constantly changing their connections, this also means that the network may be too chaotic and noisy making it too unrealistic to be useful for analysis, as connections are constantly being formed and broken.

**Watts-Strogatz** For the Watts-Strogatz model, we can expect to see a small-world network with a higher number of connections being formed and broken over time due to the high probability of connections being removed. Popular individuals may still maintain a higher number of connections, but due to the high probability of connections being broken, the gap between the number of average friends and average friends of friends for popular and non-popular individuals may not be as large as in other configurations. We can also expect to see a greater level of clustering, with similar individuals maintaining connections with each other, and non-popular individuals having a higher number of connections compared to popular individuals.
Contrary to what was expected we see a high number of popular individuals

maintaining their connections with each other this time, this results in a network made up of mostly strong connections with only a low rate of weak connections changing over time even if the probability of losing friends is much higher this time.
This configuration makes also the model slow and heavy during run time.

**Preferential Attachment** In the preferential attachment model, we can expect to see a network where new individuals are more likely to form connections with popular individuals due to their higher number. This may result in popular individuals having a higher number of connections and a larger gap between their number of average friends and average friends of friends compared to non-popular individuals.
As the number of individuals in the network increases during run time, we can see that a few highly popular individuals end up monopolizing a large number of connections, leading to a highly skewed distribution of connections within the network.

## 5 Experiments - Game Theory

In this section, we will be using the game theory model which sees individuals make decisions based on their friendliness.
We utilize Erdos(Random Graph) and Watts(Small World) models as initial configurations of the network, hence we expect to see how the initial state of the network will affect the decision-making process of the individuals.

### 5.1 Configuration 1

*num-of-nodes: 10, prob-making-friends: 0.5, prob-removing-friends: 0.5, prob-being-popular: 0.25*

**Random Graph** Using a random graph as the initial state, we can expect to see a more randomized decision-making process, as individuals may not have strong connections and may base their decision on random choices. Running the model, we can observe that individuals do tend to make more randomized decisions, with some friendly individuals initiating links with less friendly ones, this means that the network tends to be less stable as friendly and unfriendly individuals may form weak connections between them.
Finally, we can see this happening also in the histograms as the averages tend to vary greatly for each individual, popular or not.

**Small-World** Using a small world as the initial state, we expect the model to maintain the initial connections as they may be stronger, while also seeing

popular individuals maintaining more connections compared to a normal Watts-Strogatz model. Running the model, we can observe that there are some instances where individuals do tend to consider their connections more in their decision-making, with some connections being stable throughout the simulation, but we can also see a lot of randomization in the model.
Finally, we can observe that in some instances the averages tend to stay around a certain range since friendlier individuals and also popular ones tend to maintain certain connections between them, but some other instances have the histograms varying by a high margin.

### 5.2 Configuration 2

*num-of-nodes: 40, prob-making-friends: 0.7, prob-removing-friends: 0.3, prob-being-popular: 0.15*

**Random Graph** This time with a higher probability of making friends, we can observe that the simulation tends to be less random and chaotic, even with a greater number of individuals.
Even if there is a low number of popular nodes, individuals tend to be friendlier and form strong connections between themselves, with a low number of weak connections. Histograms also tend to stay within a certain range, letting us observe a more stable simulation without the danger of creating a fully connected network.

**Small-World** Using Small-world as an initial configuration we can see similar results, as with a random graph, the only difference is that there is an equivalent number of connections between popular and non-popular individuals.
On the other hand, we can observe a more evolving network where popular individuals still maintain a higher number of stronger connections than non-popular ones. The higher number of individuals, makes connections important in decision-making, with individuals prioritizing maintaining strong connections and considering forming more connections with popular individuals.
Looking at the histograms we can see that this difference is not as pronounced as in other models, with a small difference in the averages.

### 5.3 Configuration 3

*num-of-nodes: 60, prob-making-friends: 0.3, prob-removing-friends: 0.7, prob-being-popular: 0.6*

**Random Graph** With this configuration, we have a higher probability of removing friends than making new ones. We can expect to see a more isolated group of individuals, with fewer connections between them. Running the model, we can observe that individuals tend to be more selective in their decision-making, considering the strength of their current connections before deciding to remove or add new ones.
Having said that, the fact that there is a higher probability of losing friends means that the network will still have a pronounced randomization in its connections. Friendly individuals may choose to maintain their connections with other friendly individuals, while non-friendly ones may decide to isolate themselves more than with other configurations.
We can also see this in the histograms, with the means tending to be lower than before, indicating a lower average number of connections.

**Small-World** Using a small world as the initial state, we can expect to see a similar trend. However, the influence of popularity may be more prominent in this case, as popular individuals may have a higher initial weight and be more likely to maintain their connections.
We can observe that the network tends to be more stable, with fewer changes in the number of connections over time.
The histograms also show a high number of connections between individuals, with spikes in some popular individuals.

## 6 Conclusions

Based on the results of the simulations described in this paper, it is difficult to identify a single model that is consistently better for simulating this kind of friendship network. Each of the models showed a unique characteristic that may make it more suitable to isolate certain scenarios the Erdos-Renyi model is useful for understanding the formation of friendships in large, diverse populations where there are no clear preferences or biases. However, the model does not provide insight into the evolution of friendship networks over time, as the network tends to become fully connected rather quickly. The Watts-Strogatz model is useful for understanding the formation of friendships in small, homogeneous communities where there are strong ties among neighbours, with also some level of exploration and diversification. The model generates networks with high clustering and small average path length, which are characteristic of small-world networks. However, the model may not be as suitable for studying the impact of popularity on friendship networks, as the results show that non-popular individuals tend to have a higher average number of friends compared to popular individuals thanks to their similarity. The Barabasi model is useful for studying networks where a few individuals have many connections and the majority of individuals have only a few connections. The model generates scale-free networks, which are characteristic of many real-world networks. The model also provides

insight into the impact of popularity on friendship networks, with popular individuals tending to have a higher average number of friends compared to their average number of friends of friends. Finally, the game theory model is useful for studying the impact of strategic behaviour on the formation of friendship networks. The model generates networks that tend to become more centralized over time, with a small number of highly connected individuals and a large number of poorly connected individuals. On the other hand, this model is too simplistic, relying on only the friendliness of individuals and their popularity may not be enough to have a deep insight into a simulation of this kind, with more time and study we can make individuals have more meaningful characteristics while also rewarding connections between similar ones, while also trying to make friendlier people include non-friendly ones, as this is usually the case in a group of people, as it is now, the model is also prone to randomness which should not be the case in a game theory model.
It is important to consider the specific goals and characteristics of the friendship network being studied when selecting a simulation model. Different models may be more suitable for different scenarios, and it may be necessary to consider multiple models to fully understand the complex dynamics of friendship networks. Overall, these simulations showed that popular individuals tend usually to have a larger number of friends compared to non-popular ones, where they have a higher number of average friends of friends, hence making popular individuals consider themselves adequate individuals.

## A How to run the models

Running the models is really easy:

- First configure the parameters you want to have during setup and run time, by operating the sliders.
- Once the parameters have been chosen, select the desired model to start the network with, there is a button for each of the available starting models, remember that the game-theory model is not compatible with the initial Barabàsi-Albert setup, this will set up the turtles in a pre-configured network based on the previously chosen parameters, also a small histogram will show which turtles are popular and which not ( 0.5 not popular, 1 popular).
- Finally, select which model to run, by clicking on the available buttons.
- The results of the simulations can be seen either in the world window or through the histograms showing the average number of friends and friends of friends for each turtle.

## References

1. Feld, Scott L. “Why Your Friends Have More Friends Than You Do.” American Journal of Sociology, vol. 96, no. 6, 1991, pp. 1464–1477. JSTOR, http://www.jstor.org/stable/2781907. Accessed 28 Dec. 2022.
2. Coleman, J. S. (1961). The adolescent society. Free Press of Glencoe.