[CS166] Network Simulation Report

March 26, 2020

Network Simulation
Report

Taha Bouhoun

Contents

1 Introduction

Building simulations using adaptive networks offers us the flexibility to model not only the attribute of nodes but also the relational data between the nodes. Social adaptive networks exhibit two changes:

- Dynamics on networks: illustrating the process of the change in the status of the network (in terms of opinion value attributed to each person in a social network model).
- Dynamics of networks: depicting the evolution of the network's structure (in terms of relationships between people and how social ties change over time).

This report builds upon the mentioned dynamics to model the changes in social networks by generating a random small-world network (Watts-Strogatz model and Barabási-Albert model). The process starts by attributing polarized values (which represent opinions on various topics), then randomly link nodes following the properties of small-world models (local clusters, few degrees of separation). Lastly, using update rules, I present an analysis of the resulting changes and how they relate to the theoretical analysis based on the interaction between two people (nodes).

2 The basic model

As a template for social network modeling, a randomly generated Watts-Strogatz small-world networks with parameters `(nodes=50, neighbors=5, rewiring_probability=0.5)` was assigned 50% weight to all its edges as well as uniform random attribution of either 1 or 0 to its nodes (representing opinions). Upon randomly selecting an edge, the initial network models the interactions between nodes based on the following update rules:

1. The change in opinion for two interacting nodes is:

$$\Delta o_i = \alpha \times w_{ij} \left(o_j - o_i \right)$$

   where $\alpha \in (0,\, 0.5]$ stands for the degree of stubbornness (i.e., the closest to 0 the more stubborn, the closest to 0.5 the faster opinions are swayed)

2. The ties between people reflect the absolute difference between their opinions which is expressed by:

$$\Delta w_{ij} = \beta \times w_{ij} \left(1 - w_{ij} \right) \left(1 - \gamma \mid o_i - o_j \mid \right)$$

   where $\beta \in (0,\, 1)$ stands for rate at which relationship strength changes and the coefficient γ refers to the amplitude of at which the difference in opinion impacts the weight of the relationship. In other words, if $0 < \gamma \leq 1$ then it mitigates the effect of the difference in opinion on edge weight. However, a value of $\gamma > 1$ leads the difference in opinions to decrease the weight of the relationship particularly if it exceeds the threshold:

$$\mid o_i - o_j \mid > \gamma^{-1}$$

3. Generating new relationships: The model carries a 1% probability for a new edge (weighted 0.5) to be constructed between two randomly selected nodes.

3 Proposed modifications

As an extension to the existing model, the report outlines various rules that would make the model much more realistic based on the intuition of how social interactions occur between people. Listed below the set of modification along with the motivation on how they improve the initial model:

- **Multiple topics:** Adding other dimensionalities to people's opinion would help capture the notion of compromising when forming relationships. In other words, it might not be realistic to assume that we all converge towards the same edge of the opinion's spectrum because we exhibit a rich diversity in our thoughts. However, we tend to be close to people that share similar ideas on the majority of our personal opinions. To illustrate this idea, I picked three topics associated with each node:

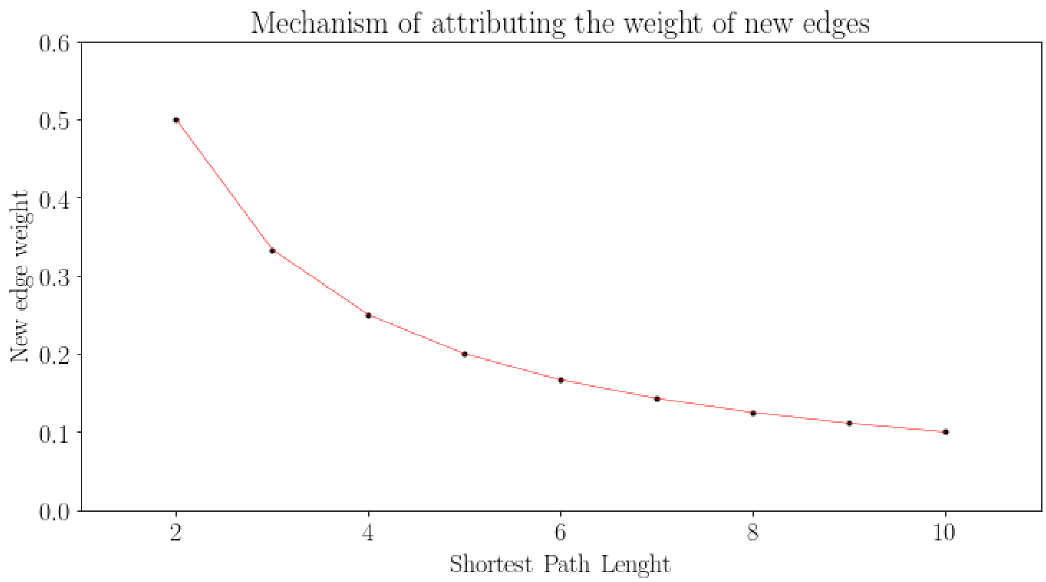
  - Political views: which are central to our lives as we tend to side with people that have the same political opinion as ours. This concept is known as tribalism as we conform to attitudes that stem from strong loyalty to one's own social group.
  - Religiousness: although religions take different forms, they share the same sense of spirituality. In the context of the model, religiousness depicts how devoted a person can be to a religion or no devotion (atheism).
  - Mindset: either a contrarian or a conformist in the sense that people either abide by social norms or they revolt against them.

  The above-listed topics are randomly selected as a subject of conversation. My intuition is that different types of clusters can emerge when introducing such dimensions (e.g., people with the same opinion on three topics, then others that are similar in 2 topics, or even people who are neutral in all three topics).

- **New relationships:** The process of creating new relationships is still random (similar to the initial model). However, the value attributed to the weight of the new relationship is proportional to the shortest path length between the two nodes as opposed to using 0.5 weight. The intuition is that if two nodes are separated with only one node (friends of friends), then the initial weight for that relationship has to be stronger than a new relationship between two nodes with three degrees of separation. The reasoning is that friends of friends are somewhat credible, and there's a higher chance for them to be introduced to each other through a friend. The formula is expressed as

$$\text{Weight}\,(i,j) = \frac{1}{\text{Shortest Path Length}\,(i,j)}$$

  The corner case is when the two nodes are not connected through a path (i.e., they belong to two disconnected networks), in such a case, the degree of separation is calculated as the diameter of the Largest (Giant) Component of in the whole network. The following figure illustrates the weights ranging from 0.5 for nodes with one node apart, up until `1 / max(shortest_path_length)`, which depends on the structure of the network.



- **Persuasiveness:** Using the TrueSkill model, each conversation is considered a game, and the outcome is measured by how close each person was swayed to the other's opinion (If person A convinced B then B is attracting to A more than A draws their opinion to B). The original model assumes that $\Delta o_i = -\Delta o_j$, which means that both people would have the same shift in their opinions but in opposite directions. To convey persuasiveness, I created a mechanism such that each person -based on their skills- shifts their opinion. Instead of converging by the same value, a random performance is drawn from each person's persuasiveness distribution (which are normally distributed) then based on the sampled performance, the highly persuasive person converges less to the other person. Mathematically, instead of having $\mid \Delta o_i \mid = \mid \Delta o_j \mid$ the change in opinion would be expressed as:

$$\text{Perfo}_i \sim N\left(\mu_i, \sigma_i^2\right)$$

$$\text{Perfo}_j \sim N\left(\mu_j, \sigma_j^2\right)$$

$$\text{Normalized Perfo}_i = \frac{\text{Perfo}_i}{\text{Perfo}_i + \text{Perfo}_i}$$

$$\text{Normalized Perfo}_j = 1 - \text{Normalized Perfo}_i$$

$$\text{Normalized Perfo}_j \times \Delta o_i \neq \text{Normalized Perfo}_i \times \Delta o_j$$

Example: Taking a conversation between two people with different persuasiveness skills.

```
Opinion on a topic:
Person i:  1
Person j:  0

Persuasiveness skill:
Person i:  trueskill.Rating(mu=8.000, sigma=1.000)
Person j:  trueskill.Rating(mu=5.000, sigma=1.000)
```

```
Sampled persuasiveness performance:
Person i:  7.4126726998635935
Person j:  4.458937620856932

Normalized performance:
Person i:  0.6244033033097144
Person j:  0.3755966966902856

Opinion on a topic:
Person i:  0.9943660495496457
Person j:  0.009366049549645716

Updated persuasiveness skill:
Person i:  trueskill.Rating(mu=7.796, sigma=0.993)
Person j:  trueskill.Rating(mu=5.204, sigma=0.993)
```

We notice that person_j has less persuasiveness skill compared to person_i and the drawn performances reflect that difference. After normalizing, the opinions of both people are updated by multiplying by α and the weight of the relationship. Given that person_i was more convincing, their opinion was less swayed away from the original value (-0.00563). However, the person_j was less persuasive hence their opinion shifted more from its original value (+0.00936)

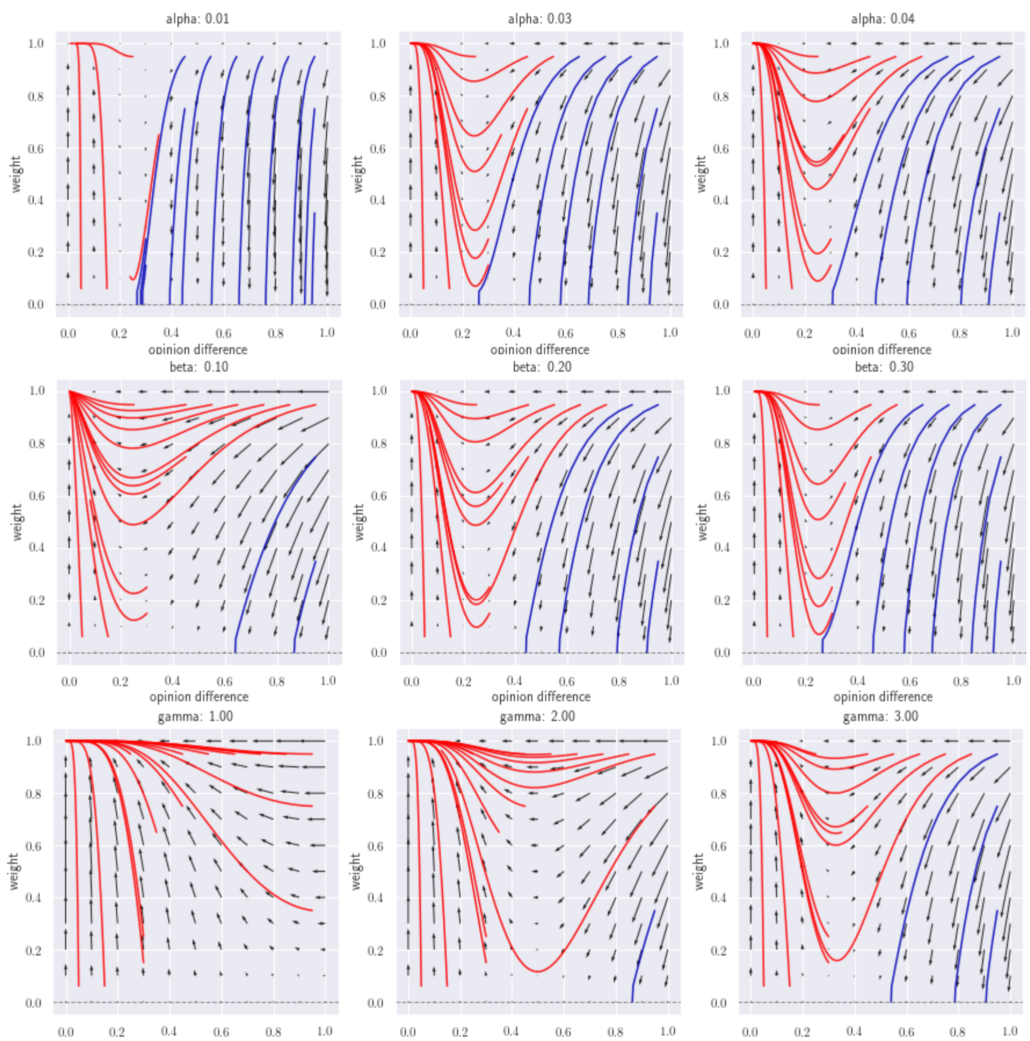
3.1 Model assumptions

- All people have the same level of socializing (ignoring extroversion and introversion)
- Conversations happen only between two people (no group discussions).
- The topics are represented as an array of values between 0 and 1. When initializing the network, each node is given a value of either 1 or 0 for each of the multiple topics added to the simulation. Thus, we assume that all topics are one-dimensional (meaning that all topics are polarized).
- The update is asynchronous, which means that interactions don't take place at the same time step. Although it might be more realistic to assume that multiple conversations between people are happening a the same time. Both approaches would result in the same patterns, given that a person can only engage in one conversation.
- The strength of the new relationship depends on the degree of separation between the nodes. Although realistic, one can make an argument that due to social interconnectedness, people can build strong relationships based on first impressions even if they're multiple degrees of separation away.

4 Local analysis

- **Multiple topics:** Similar to the basic model, the opinion of two people on a given topic changes as they discuss it. The difference is that the topics are independent (i.e., People can differ significantly in two topics, but if they randomly select the third topic that they agree on, their relationship will strengthen). On the other hand, we can expect that nodes that differ in opinion on two topics would have a low chance of rapprochement, and the edge weight linking between them would witness a decrease (2/3) of the time.

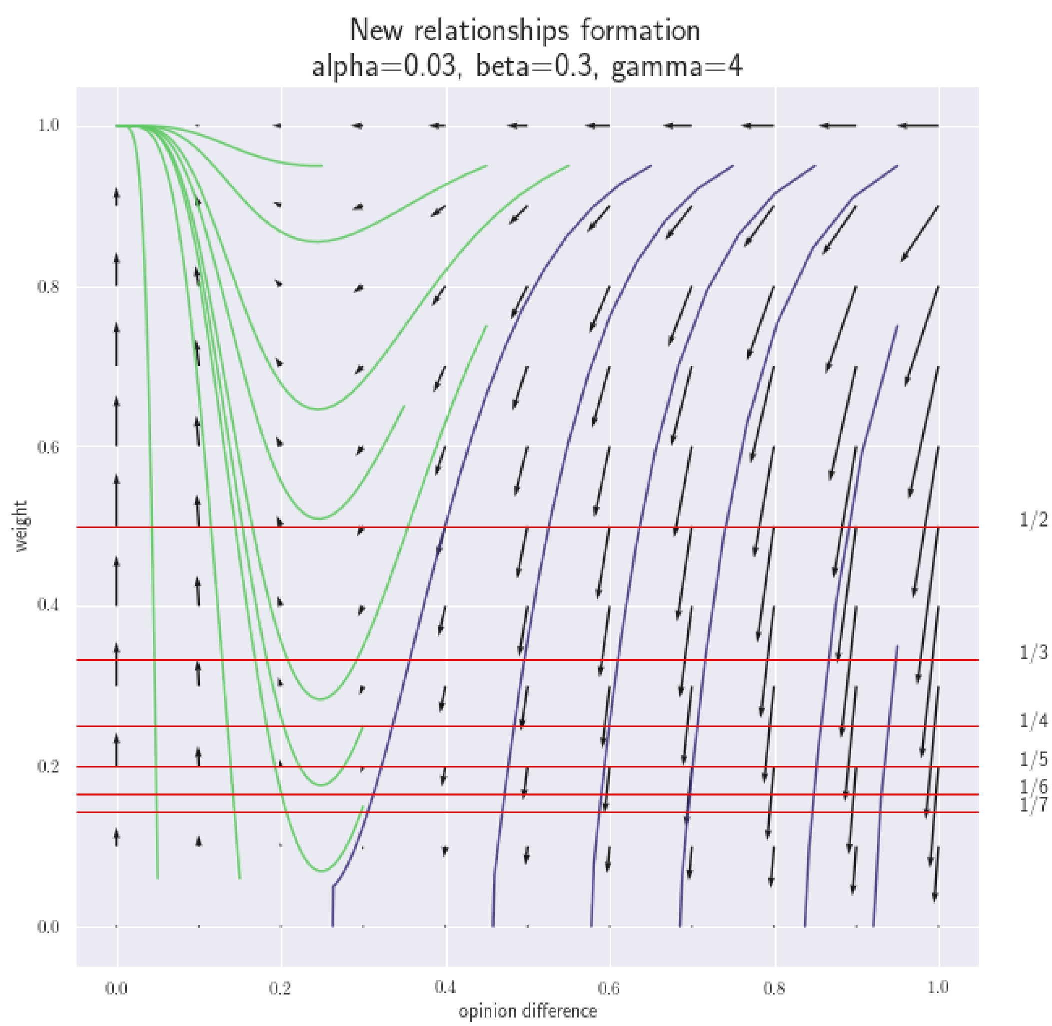
  The vector fields below show the potential path for a given pair of nodes based on their difference in opinion and their edge weight. The values of α, β, and γ are varied across three values (more values are added on the appendix).



  The vectors highlight the intensity/direction of the change in both opinion and weight for a pair of nodes. The red lines illustrate the paths of convergence/closing-up and the blue line mark the paths of divergence/cut-off. In the context of multiple topics, these forces might repel each other if the pair of node was picked twice on different topic (one in which they

agree on and one which they don't). Regardless, the overall pattern would still obey the same patterns of convergence and break-ups.
Controlling for the default values $\beta = 0.3\,\gamma = 4$, an α value of 0.1 makes a pair of node with opinion difference of 30% or more and edge weight of less than 0.5 break-up in the long run if they keep interacting with each other. Upon driving the value of α up to 0.4, the break-up only occurs if the difference in opinion is at least 50% with an edge weight 0.5. The same analysis can be drawn for β and γ by controlling for the two other variables. To make the simulation interesting, I picked the values that lead roughly to an equally shared area of being prone to converge and diverge (i.e., values that has slit the area into half between red and blue paths). These values turns out to be the default values.

- **New relationships:** Fixing the values of the parameters at their default values $\alpha = 0.03$, $\beta = 0.3\,\gamma = 4$ we can draw the potential line at which the new relationships can be formed. Each red line stands for a newly created edge weight between two nodes. Notice that for a 0.5 weighted relationship, the gap in the opinion of more than 40% leads the nodes to break up. Although, the worst case is when two newly created nodes are 7 degrees of separation away (weighted 1/7), there has to be a 30% difference in opinion for that relationship to break.



- **Dunbar Number:** The cognitive ability of humans suggests that there's a limit of how many stable social relationships we can maintain. Anthropologist Robin Dunbar suggests that the number lies between 100-200. Although this simulation doesn't deal with large numbers of nodes, I attempted to test whether the sum of edge weights for each node at the end of the simulation is stable around a number. (*See function* `dunbar()` *in code source*)

- **Friendship paradox:** A phenomenon first observed by the sociologist Scott L. Feld, which states that -on average- most people have fewer friends than their friends have. Through computing the average degree of neighbors, I test to what extent running the simulation feeds into/relaxes the friendship paradox.

- **Small-World test:** Since I'm modeling human social networks and their changes based on interactions, it is vital to check whether the network maintains the Small-World properties after running the simulation. The properties are high local clusters and right-skewed distribution of shortest paths. The parameters used for testing the properties are σ and ω (*See function* `test_small_world()` *in code source for details*)

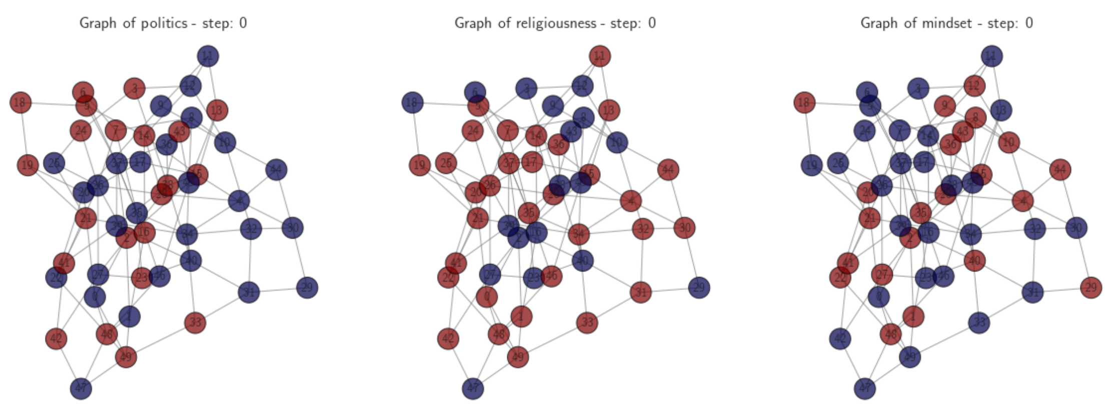
5 Implementation and Analysis

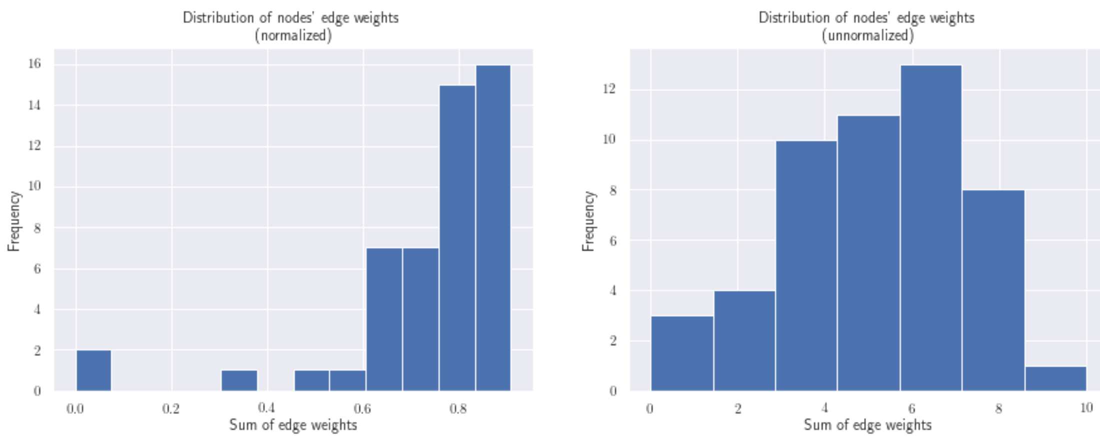
The proposed modifications (Multiple topics, New relationships) are applied to two examples of small-world networks after controlling for the number of nodes, number of edges, and the average degree of the graph. The results below show the long term changes given the update rules. The third model adds the Persuasiveness using TrueSkill on the Watts-Strogatz graph. The color coding map is *seismic* (red, blue) $\rightarrow (1, 0)$

5.1 Watts-Strogatz model:

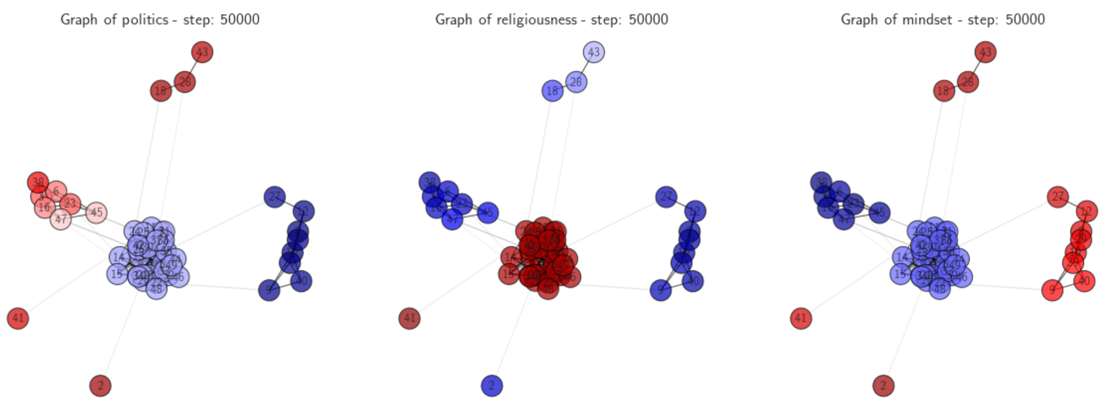
Watts-Strogatz network takes the number of nodes, the k-nearest , and the rewiring probability as parameters. The network is constructed in a ring shape topology, then each node would have one of its edges randomly rewired to another node. The initial layout for the three topics are:

```
Number of nodes:  50
Number of edges:  100
Average degree:  4.0000
Average degree of neighbors in:  4.45
```

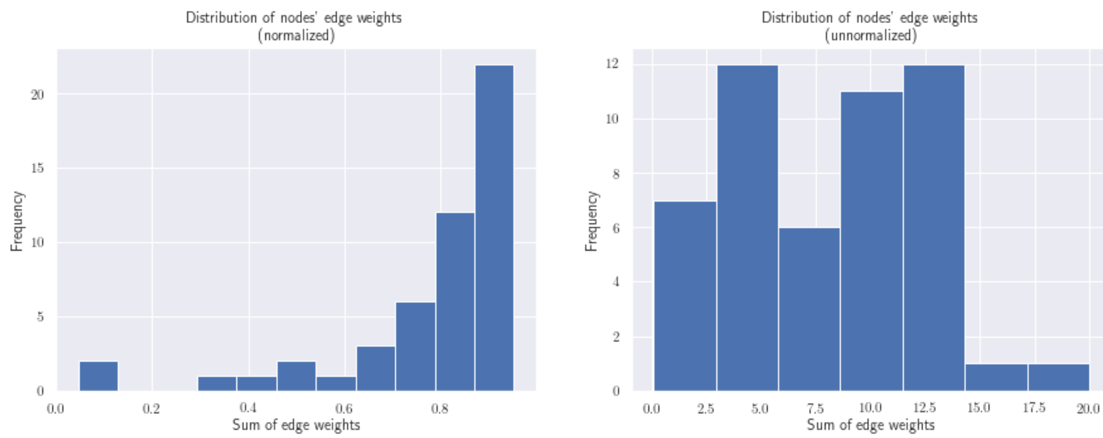




The final configuration of the network shows various local clusters that contain nodes that share roughly the same opinion on all three topics. As opposed to the original model (where it converges towards only two clusters), adding more topics seem to create different variations of clusters. Intuitively, the number of these clusters scales according to the number of topics that we attach to each node. In other words, one topic leads to two clusters, two topics leads to four clusters ([1,1],[0,0],[1,0],[0,1]), and finally sharing three topics would amount to eight possible combinations of clusters. Mathematically, we can write it as: `clusters =` 2^{topics}




```
Number of nodes:  50
Number of edges:  207
Average degree:  8.2800
Average degree of neighbors in:  10.97
```



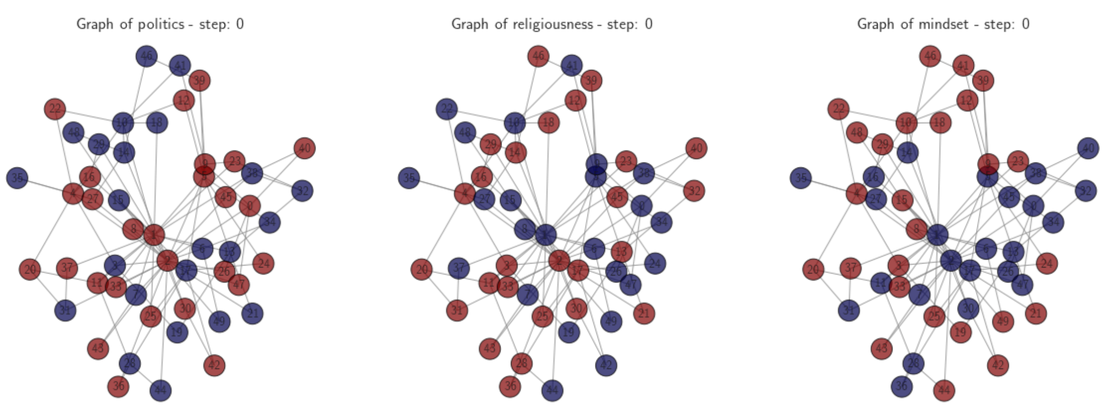
The normalized Dunbar distribution barely changed compared to the initial distribution. However, the average degree of the graph and the average degree of neighbors doubled (from 4.00 and 4.45 to 8.28 and 10.27 respectively) suggesting that the friendship paradox is not apparent in this graph. Note that also the number of edges in the graph doubled from 100 to 207 edges, and the resulted graph passed the Small-World test by the following values:

```
Small-world:  True sigma = 1.22 > 1
Small-world:  True -1 < omega = -0.05 < 1
```

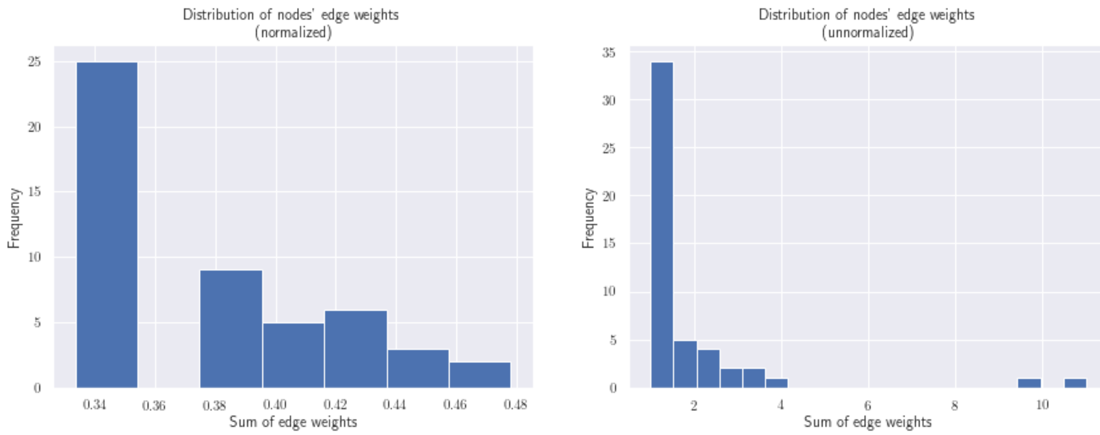
5.2 Barabási-Albert model:

Barabási-Albert network uses the concept of preferential attachment to give rise to local clusters within the network. It takes as arguments the number of nodes and the number of edges to be attached to newly added nodes. The process is to add one node at the time and associate higher probability for new nodes to connect with existing high-degree nodes.[1]
Given the preferential attachment in the Barabási-Albert model, the distribution of the "Dunbar" number is skewed to the left because few nodes are carried high degree relative to the rest of the nodes which makes the friendship paradox -on average- prevails in this graph

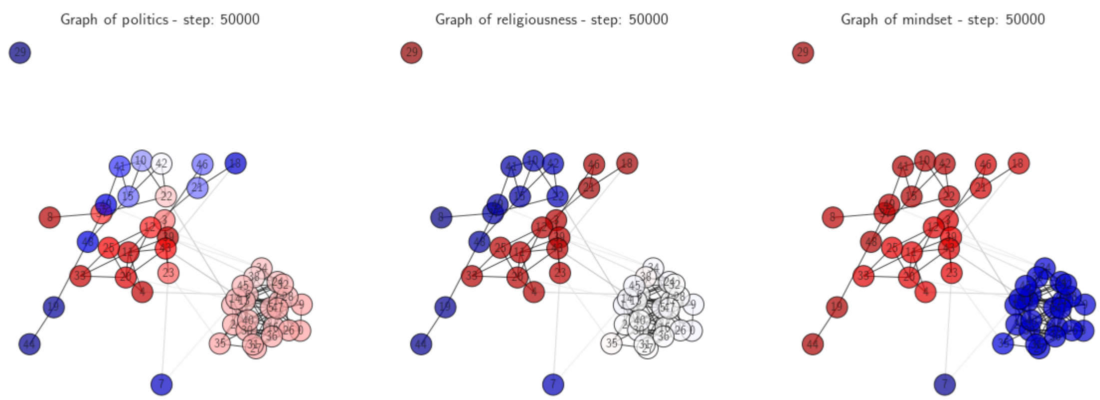
```
Number of nodes:  50
Number of edges:  96
Average degree:  3.8400
Average degree of neighbors in:  6.16
```



[1] *It is essential to mention that although the average degree of the previous model is approximately the same as its average degree of neighbors, Barabási-Albert model has an average degree of neighbors that is 50% greater than its average degree.*



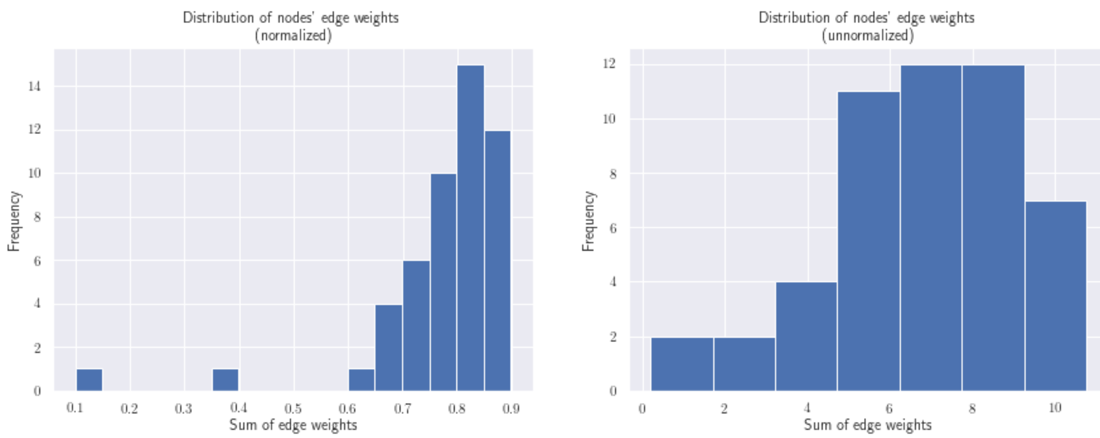
After running the simulation, the network is segregated into two giant clusters (and a single node on the top left). Strong clusters share the same opinion on all three topics (bottom-right), but other clusters also form by sharing two out of the three topics. but they're not as tight as the former. The skewness of the normalized Dunbar distribution is shifted to the right as opposed to the initial configuration, which entails that nodes are strongly connected to each other compared to the initial setting. Similar to the final configuration on the Watts-Strogatz model, the friendship paradox in Barabási-Albert model is not maintained since the average degree of the graph and the average degree of neighbors roughly the same.




```
Number of nodes:  50
Number of edges:  187
Average degree:  7.4800
Average degree of neighbors in:  8.26
```

The resulted graph passed the Small-World test by the following values:

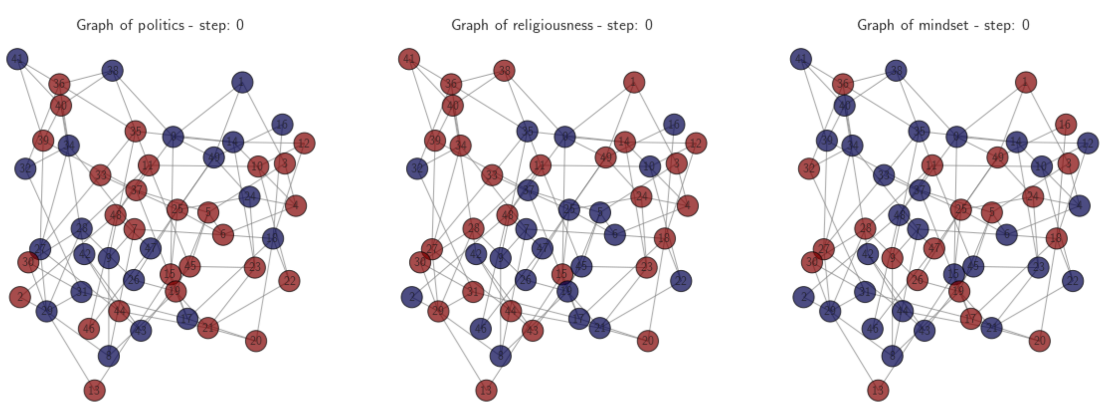
```
Small-world:  True sigma = 1.63 > 1
Small-world:  True -1 < omega = 0.47 < 1
```

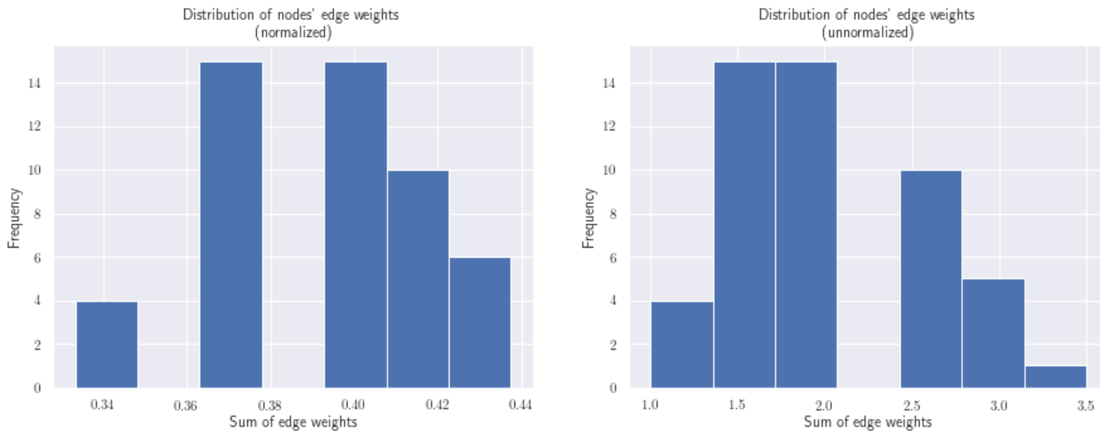


5.3 TrueSkill for persuasiveness:

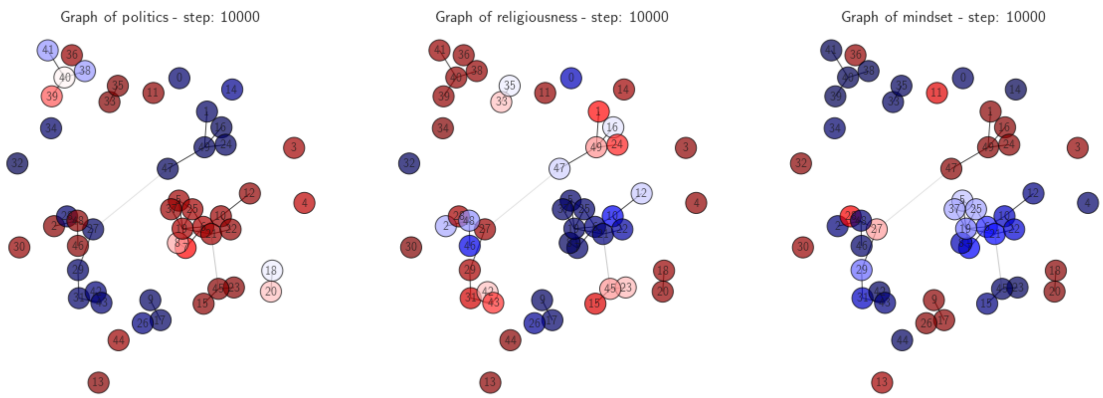
As described in section 3 (*Proposed modifications*), persuasiveness is gauged based on an initial distribution of skill for all nodes. As a result, a conversation can be considered as a match between two battling opinions, and the winner is the person that sways the other person further from their opinion. The nodes are initialized with a normal distribution for their persuasiveness skill centered at values ranging from μ=5 to μ=10 to conveying that people have different levels of skills (distributions are attributed the same $\sigma = 1$). Using the Watts-Strogatz model with the same parameters as (5.1), we get the initial configuration.

```
Number of nodes:  50
Number of edges:  100
Average degree:  4.0000
Average degree of neighbors in:  4.35
```

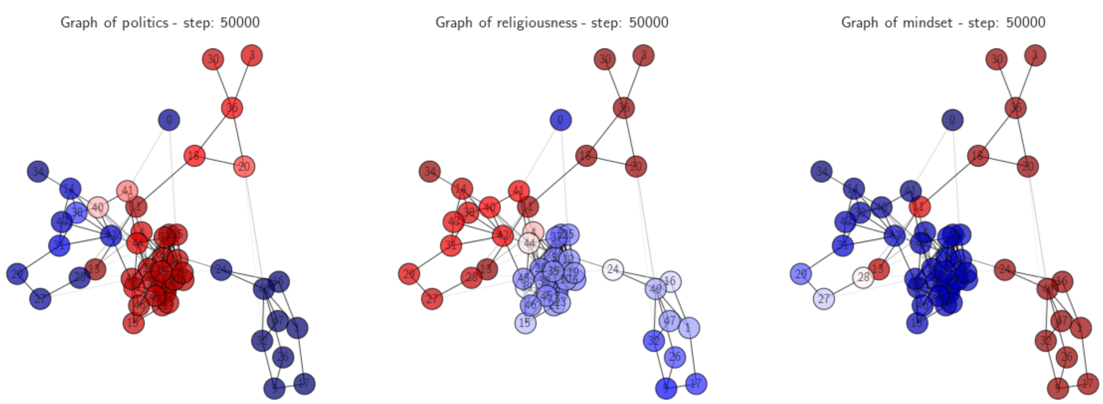


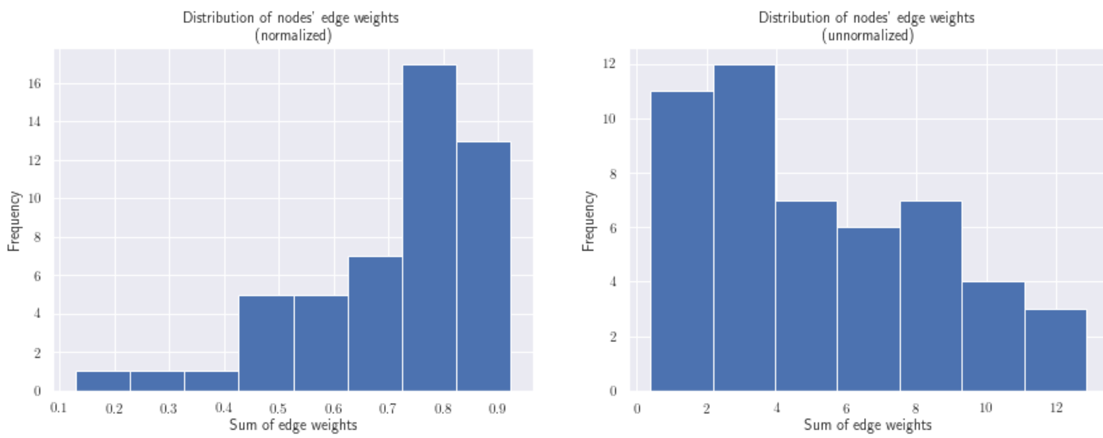


After 10,000 time-steps, nodes seem to loose their edges (*yes pun intended*) and became more scattered but some of them are forming groups especially if they share the opinions in all three topics as shown in the graph below:



Upon reaching 50,000 time-steps, the nodes form clusters and the connectedness of the whole network is restored (less chance of finding stray nodes). The combination of clusters with nodes sharing the same opinion on all three topics still holds. The difference between average degree of the graph and the average degree of neighbors is significant (about 2) which suggests that the friendship paradox is less likely to occur on average.

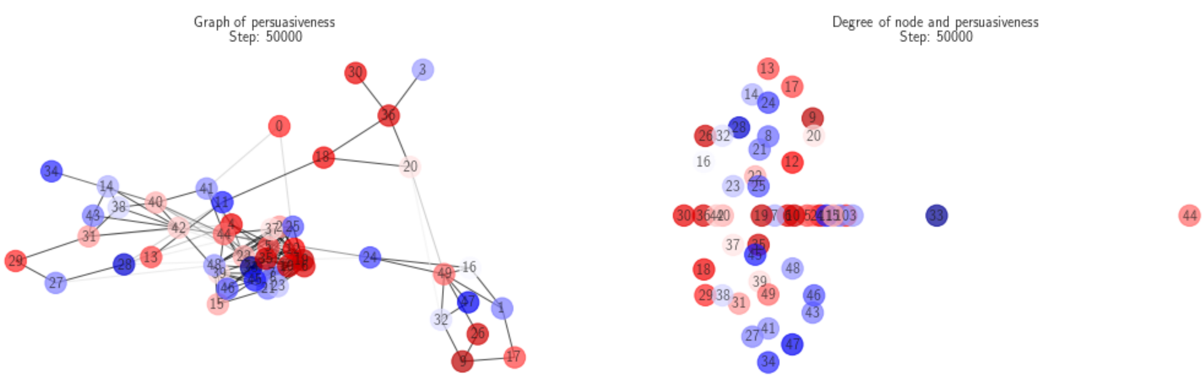





```
Number of nodes:  50
Number of edges:  144
Average degree:  5.7600
Average degree of neighbors in:  7.75

Small-world:  True sigma = 1.7 > 1
Small-world:  True -1 < omega = -0.01 < 1
```

The following graph highlight the persuasiveness mean for each node relative the the minimum and maximum persuasiveness level in the graph. It doesn't seem that nodes with high mean (red ones) are necessarily in the center of clusters as mentioned in the left graph. The right-side graph highlights the degree of the node as well as their persuasiveness score, it doesn't appear to be clear correlation between the degree of a node and its level of persuasiveness.



6 Conclusion

Social relations are a result of compromising some of the differences between us, given the number of shared interests. The simulation above showed many combinations of clusters and how the relationship between them strengthens over time when opinion is shared, and weakens with people that differ on two or more topics. The Dunbar number doesn't seem to be constant for all nodes (the sum of their edges' weights). In addition, the friendship paradox doesn't hold (with the Barabási-Albert model). The formation of clusters could suggests that people would converge to having -on average- the same number of relations.
Similar to real life social networks, clusters can vary as many interests/topics we could share. The default parameters of the model seem to manifest realistic behaviours.

Appendix

- Source Code: contains further analysis of the parameters values as well as incremented changes in the simulations (every 10,000 time-steps).
  I also attempted the models on a Facebook graph but it was computationally expensive to apply the updates mentioned in the modifications.
  The code also contains a table for the TrueSkill model and the updated results of the people's opinions and their persuasiveness' parameters.

- LO/HC Application:

  #networks [HC]: using adaptive network as a template for modeling social networks.

  #modeling [HC]: Using simple social interactions rules to construct a societal level analysis given the defined assumptions.

  #simulation [HC]: a methodology to analyze non-linear functional relations such as social relationships.

  #network_analysis [LO]: Analyzing the critical values of the parameters of the models as well as the characteristics of the network after updating its state (Clustering, Degree distribution, Small-World test)

  #network_modeling [LO]: Applying the mathematical rules of the change in weight and opinions and adding modifications that better reflect real social networks.

  #interpret_results [LO]: Interpreting the results of the simulations by drawing from the results and justifying their applicability on real-life scenario.

  #python_implementation [LO]: implementing the code for the simulation in Python

References

[1] Scheffler, C. (2020). CS166 Assignment 3: Network simulation. *Minerva Schools at KGI*

[2] Leskovec. J., (2018). Stanford Large Network Dataset Collection. *Stanford University*