THE EFFECTS OF HETEROGENEITY IN THE NUMBER OF CONTACTS

Introduction

We are going to carry out a numerical simulation of a rumor propagation process. The model I have created has the following characteristics:

- It starts with a single individual carrying a hot new.
- The population consist on 200 individuals, and each individual knows 8 other individuals.
- If two persons know each other, there is a 10% probability that they are close friends. I have added this to simulate a bipartite graph.
- The population is divided in the people that know the rumor and the people who do not (*ignorants*).
- When somebody finds out the rumor, that person can become a *spreader* (70% chance) or a *stiffer* (30%) chance. The *stiffers* never tell the rumor to anyone. The *spreaders* do, under the following conditions:
    - ✓ They can only tell the rumor to people they know.
    - ✓ They will tell the rumor to regular friends with an 80% chance. Even if at the first chance they have not told them, they can do it on subsequent chances.
    - ✓ They will always tell the rumor to close friends.
- The process ends when:
    - ✓ Everybody knows the rumor.
    - ✓ The same number of people knows the rumor two consecutive times, because that means that either everyone who could find out about the rumor already knows, or that the rumor has gone out of style.

To implement the model, I will use three different types of networks, and compare the results. To do this, I have created a script in R using the package *igraph*. For each network I have created 3 kinds of representation for the networks:

- The network is shown in the first place with all the edges, being grey edges the relations between regular friends and blue edges the relations between best friends. All nodes are yellow.
- The networks in the different times during the rumor expansion are shown using the following color code: grey nodes are ignorants, black nodes are stiffers and red nodes are spreaders. Only the edges between spreaders and ignorants are shown, always in black.
- The final state of the network is shown with all the edges, blue and gray, and the nodes colored red, grey and black.

Random network

In general, a random graph is a model network in which some specific set of parameters take fixed values, but the network is random in other respects. One of the simplest examples of a random graph is the network in which we fix only the number of vertices *n* and the number of edges *m*. This model is often referred to by its mathematical name ***G (n, m)***, and it was proposed by Erdös and Rényi in 1959.

Some properties of the random graph are straightforward to calculate: obviously the average number of edges is *m*, and the average degree is $k = 2m/n$. Unfortunately, other properties are not so easy to calculate, and most mathematical work has actually been conducted on a slightly different model that is considerably easier to handle. This model is called ***G (n, p)***. Here, we fix not the number but the probability of edges between vertices. Again, we have *n* vertices, but now we **place an edge between each distinct pair with independent probability *p***.

On *Figure 1* is shown the original network, on *Figure 2* the final state of the network, and on *Figure 3* are shown the different stages of the network.

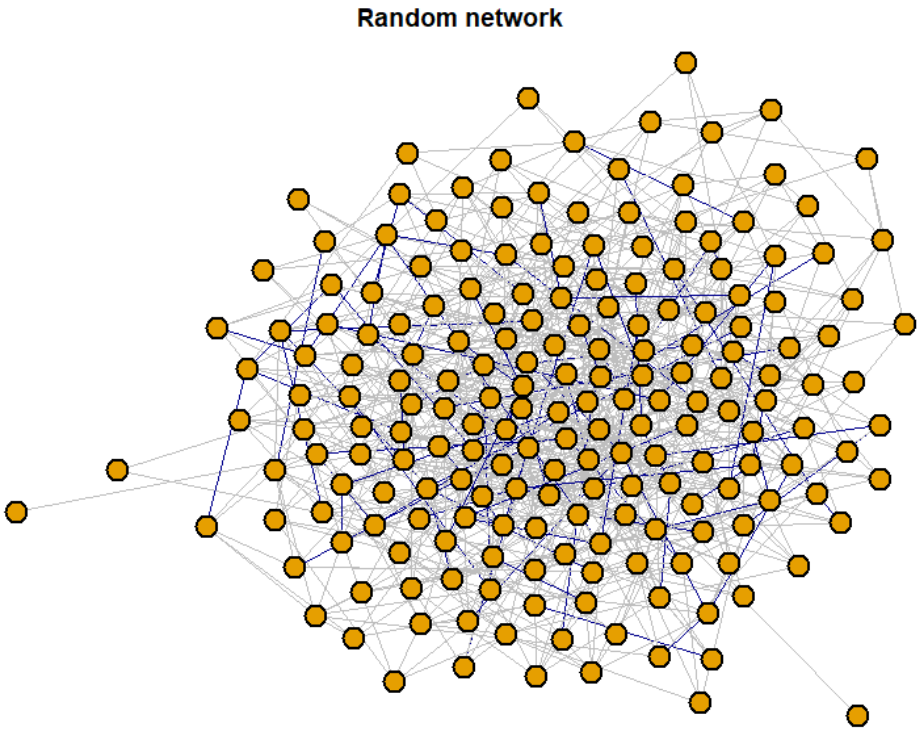

Figure 1

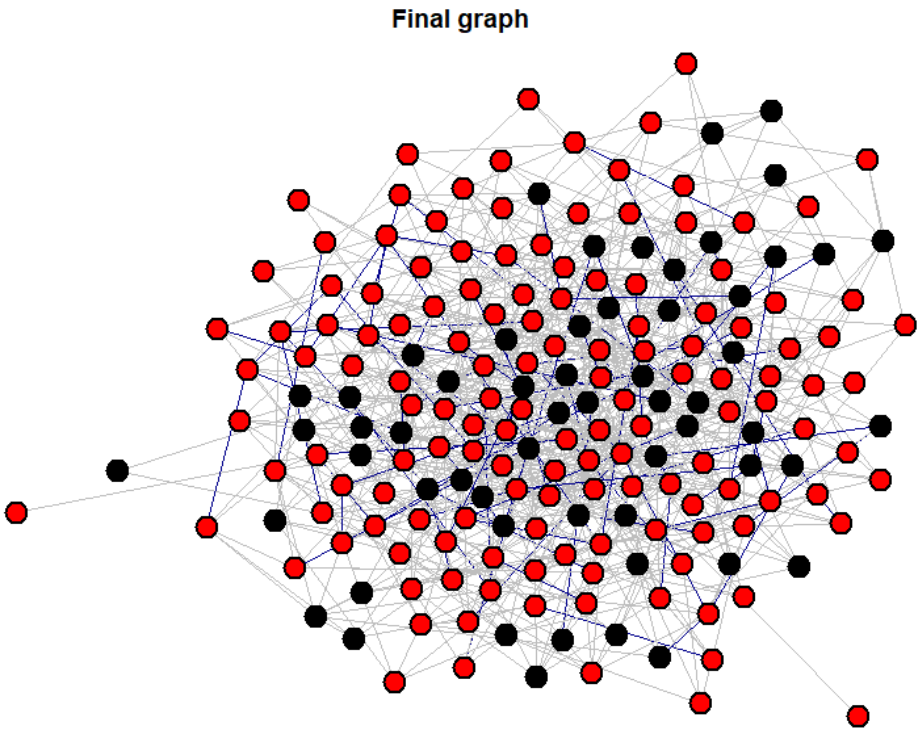

Figure 2

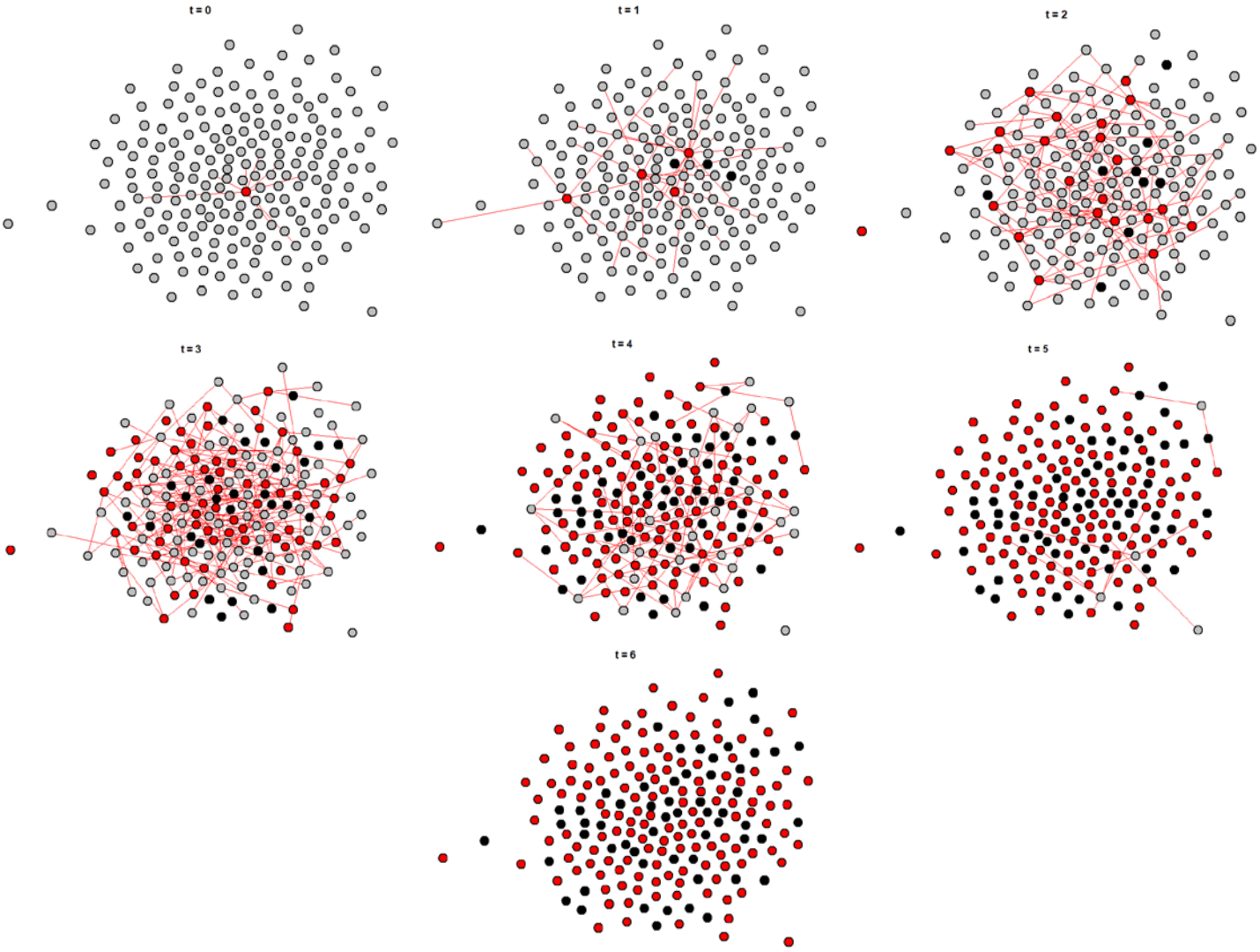

Figure 3

It has taken 6 periods for the rumor to extinguish. Everybody found out in the end. The model is completely random, we cannot see any patterns for the expansion of the rumor.

Small world network

The small-world effect is the observation that the geodesic path distance between most pairs of vertices in a network is small.

One of the least well-understood features of real-world networks is transitivity, the propensity for two neighbors of a vertex also to be neighbors of one another. Neither the random graph models nor the models of preferential attachment generate networks with any significant level of transitivity, as quantified by the clustering coefficient. The *small-world model*, proposed in 1998 by Watts and Strogatz, displays both high transitivity and short path lengths simultaneously.

The small world model, in its original form, interpolates between the circle model and the random graph by moving or rewiring edges from the circle to random positions. Starting with a circle model of n vertices in which every vertex has degree c, we go through each of the edges in turn and with some probability p we remove that edge and replace it with one that joins two vertices chosen uniformly at random. The randomly placed edges are commonly referred to as shortcuts.

On *Figure 4* is shown the original network, on *Figure 5* the final state of the network, and on *Figure 6* are shown the different stages of the network.

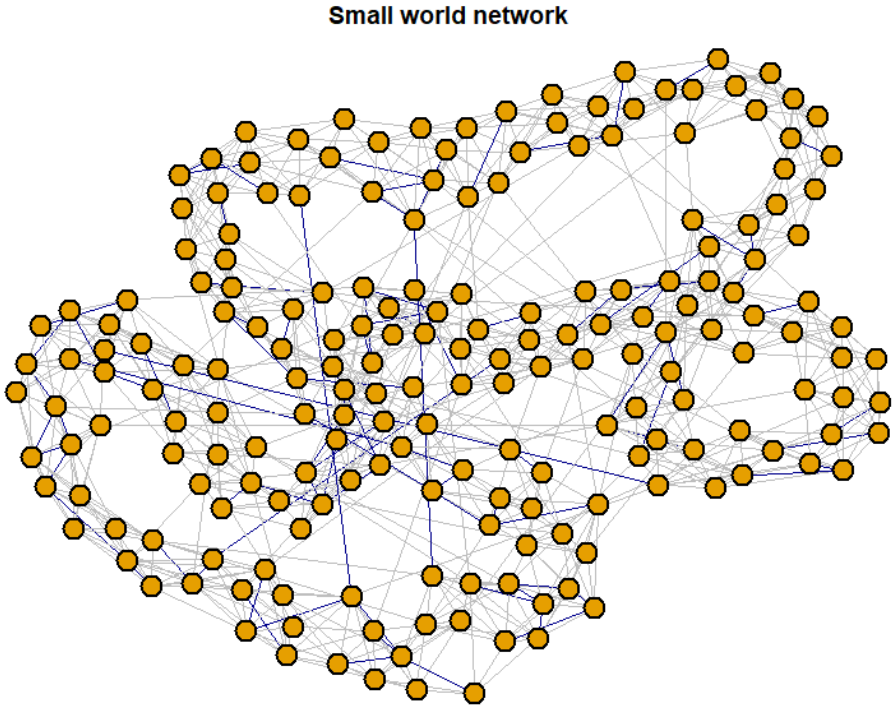

Figure 4

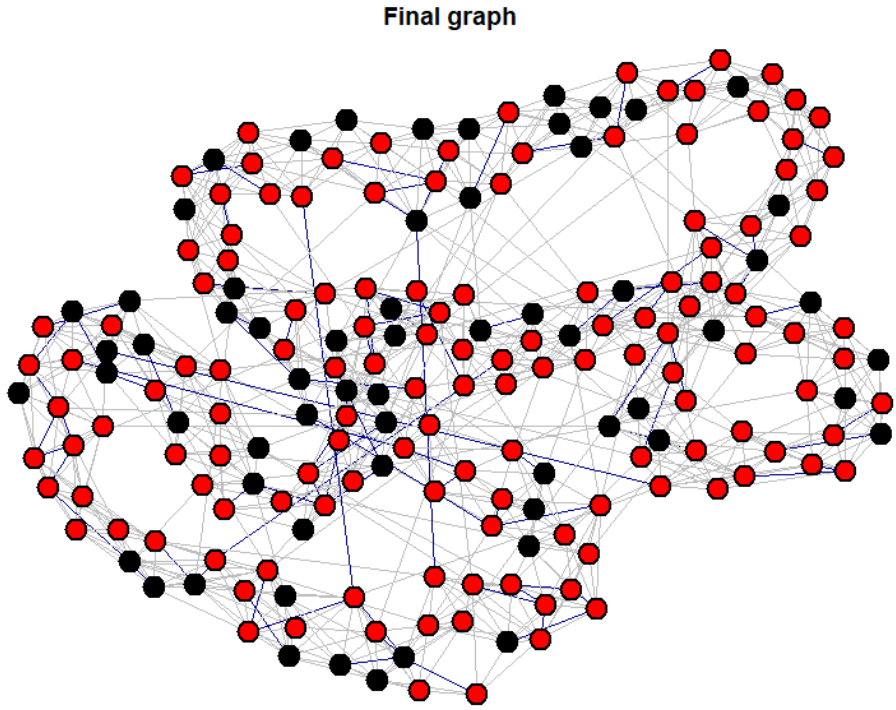

Figure 5

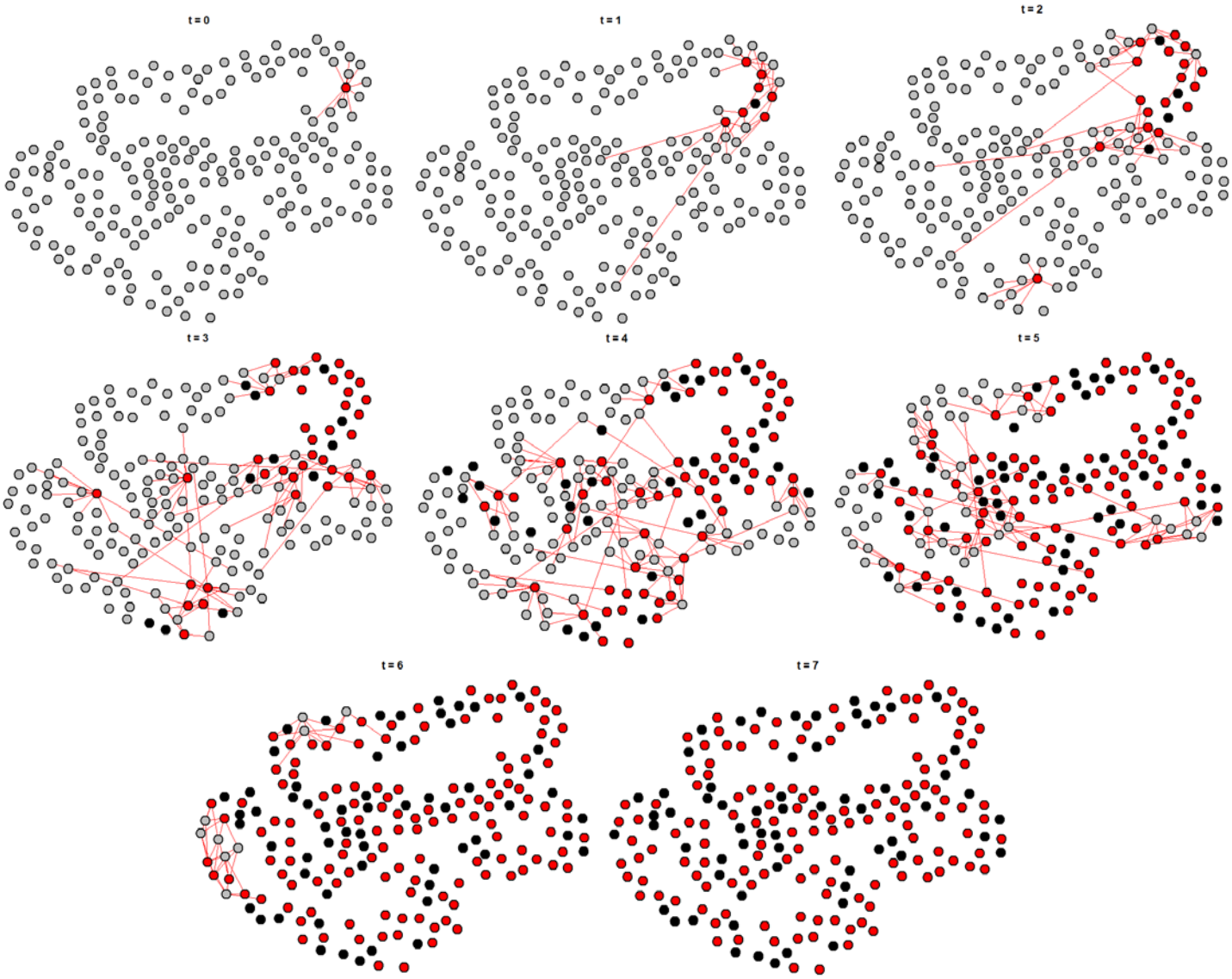

Figure 6

It has taken 7 periods for the rumor to extinguish. Everyone eventually found out. In this model, it can clearly be seen how the rumor expands by spatial closeness, except for the shorcuts that connect far nodes.

Preferential attachment network

Many networks are observed to have degree distributions that approximately follow power laws, at least in the tail of the distribution. The power law is a somewhat unusual distribution, and its occurrence in empirical data is often considered a potential indicator of interesting underlying processes. A natural question to ask, therefore, is how a network might come to have such a distribution. This question was first directly considered in the 1970s by Price, who proposed a simple and elegant model of network formation that gives rise to power-law degree distributions.

Price adapted the methods of an economist called Herbert Simon, with relatively little change, to the network context. Price gave a name to Simon's mechanism: he called it *cumulative advantage*, although it is more often known today by the name *preferential attachment*, which was coined in 1999 by Barabási and Albert.

In summary, Price's model consists of a growing network of nodes and edges in which vertices are continually added but none are ever taken away. Each node is related, on average, to *c* others (so that the mean out-degree is *c*), and the vertexes to which edges go are chosen at random with probability proportional to their in-degree plus a constant *a*.

Preferential attachment did not become widely accepted as a mechanism for generating power laws in networks until much later, in the 1990s, when it was independently discovered by Barabási and Albert, who proposed their own model of a growing network (along with the name "preferential attachment"). The Barabási-Albert, which is certainly the best known generative network model in use today, is similar to Price's, though not identical, being a model of an undirected rather than a directed network.

In this model, vertices are again added one by one to a growing network and each vertex connects to a suitably chosen set of previously existing vertices. The connections, however, are now undirected and the number of connections made by each vertex is exactly *c* (unlike Price's model, where the number of connections was required only to take an average value of *c* but might vary from step to step). Note that this implies that *c* must be an integer, since a vertex cannot have non-integer degree. Connections are made to vertices with probability precisely proportional to the vertices' current degree. Notice that there is no in- or out-degree now because the network is undirected. Connections are made simply in proportion to the (undirected) degree. As before, vertices and edges are only ever added to the network and never taken away, which means that there are no vertices with degree $k < c$. The smallest degree in the network is always $k = c$.

On *Figure 7* is shown the original network, on *Figure 8* the final state of the network, and on *Figure 9* are shown the different stages of the network.

The evolution of this network depends heavily on the starting conditions. Once it gets to a central node, the expansion is quick, bout on external nodes, it is very slow.

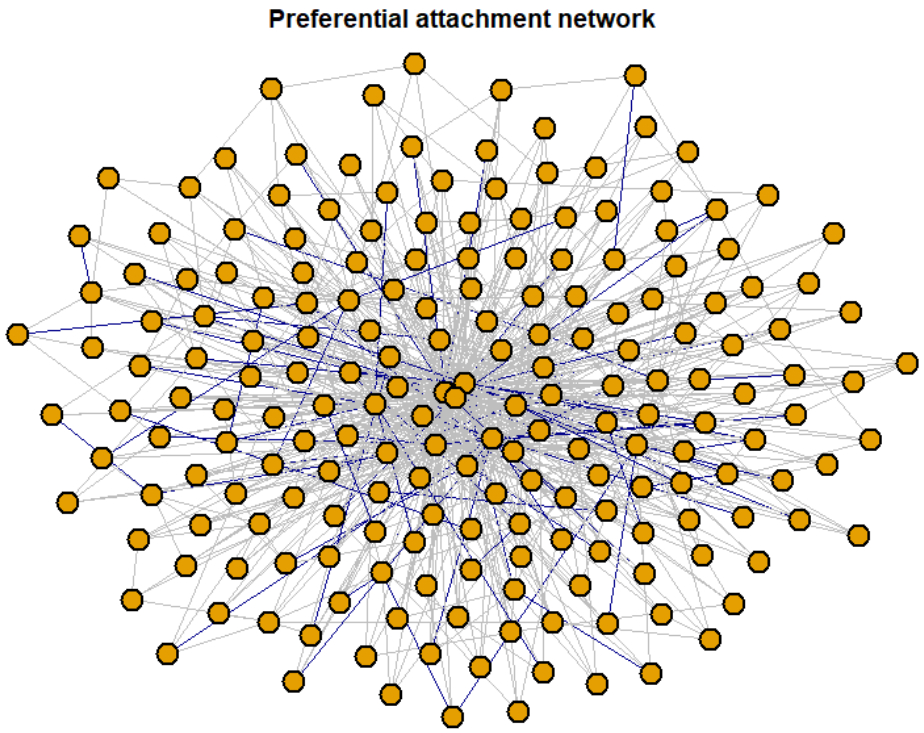

Figure 7

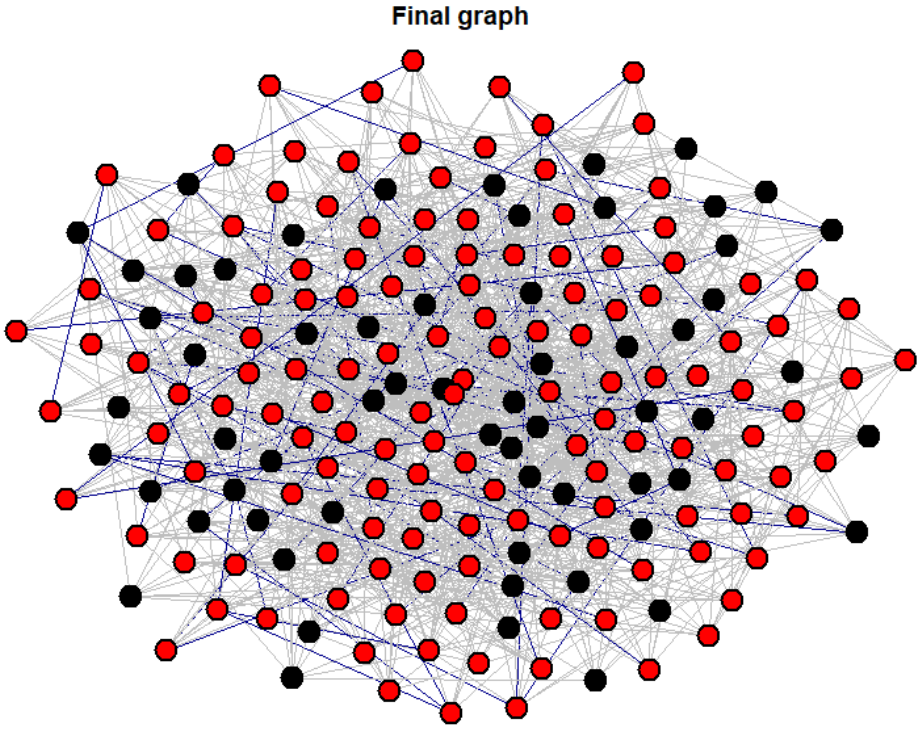

Figure 8

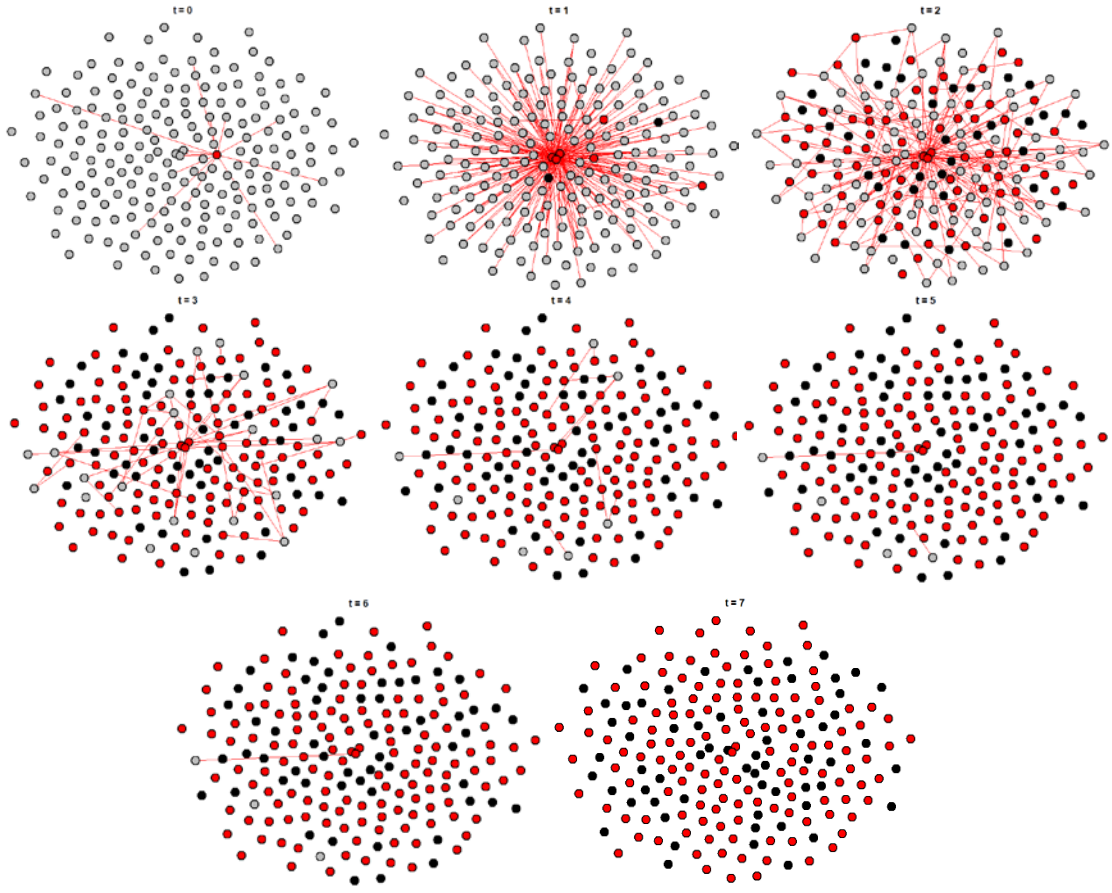

Figure 9

It has taken 7 periods for the rumor to extinguish. Everyone eventually found out. **The evolution of this network depends heavily on the starting conditions, contrary to the other two**. Once it gets to a central node, the expansion is quick, bout on external nodes, it is very slow.

Mean field comparison

On *Figure 10*, there is a comparison of the three networks (*A*: random; *B*: small-world; *C*: preferential attachment). I have calculated the mean-field solution as follows:

- The first element is 1.
- The second element is calculated as:

  $$1 + 8 * (0'9 * 0'8 + 0'1)$$

  1 is the people that already knows the secret, 8 the average number of people known, 0'9 the probability that a friend is a regular friend, 0'8 the probability that the rumor is told to a regular friend, and 0'1 the probability that a friend is a close friend.
- The rest of the elements are calculated as the second, but the second term is also multiplied by 0'7 (the probability that a person who knows the secre tells it). This is not applied in the second term, because the first person is always supposed a spreader.

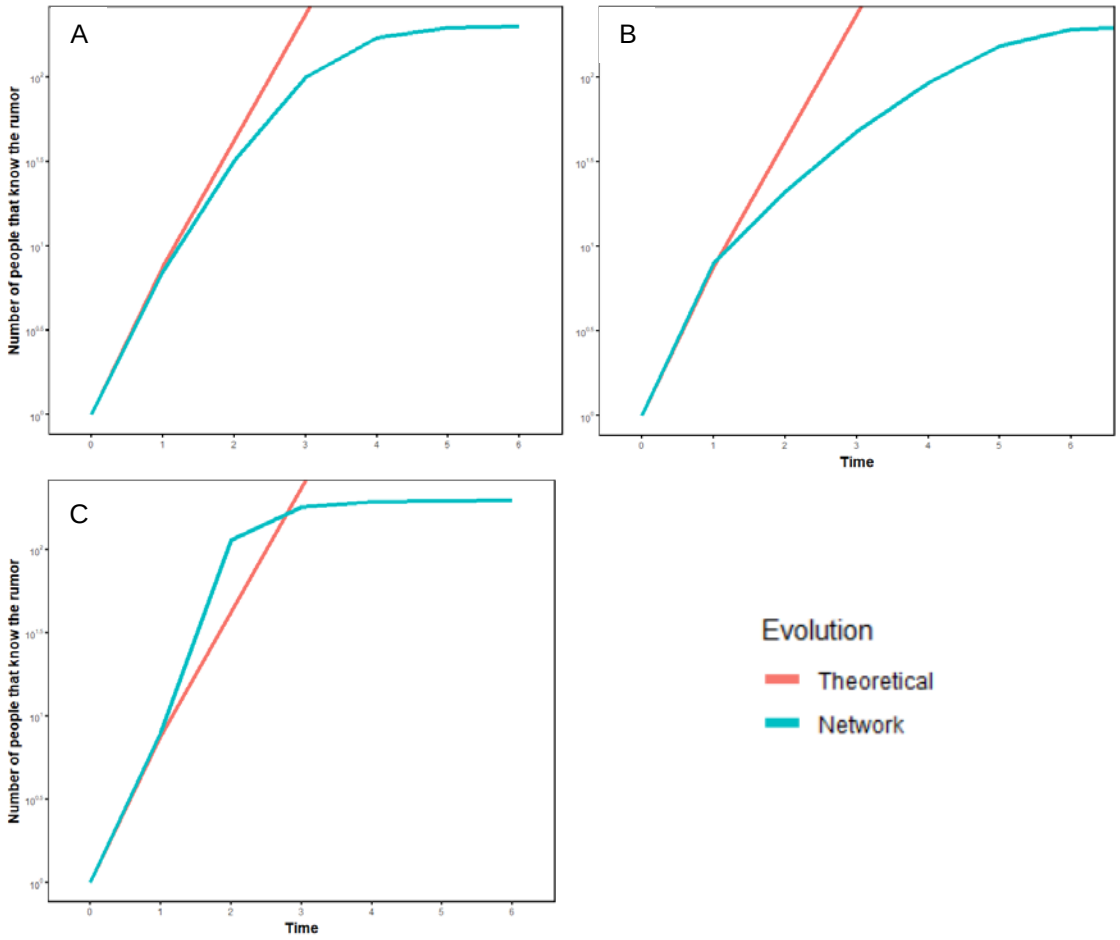

Figure 10

The small world an random networks follow similar distributions, always below the mean-field solution. The prefererential attachment network, however, has a different distribution, reaching a very high number of people in a very small time (even faster than the theoretical distribution) and then taking many generations to end. This is due to what was commented previously in the text: the central nodes expand the rumor very quickly, while the external ones are very slow. Had the rumor started on a peripherical node instead of a central one, and the number of people who knew the secret on the first generations would have been a lot less.

Eigenvector centrality

For each network, the eigenvector centrality distribution has been calculated; this is a natural extension of the simple degree centrality. We can think of degree centrality as awarding one "centrality point" for every network neighbor a vertex has. But not all neighbors are equivalent. In many circumstance, a vertex's importance in a network is increased by having connections to other vertices that are themselves important. This is the concept behind eigenvector centrality. Instead of awarding vertices just one point for each neighbor, eigenvector centrality gives each vertex a score proportional to the sum of the scores of its neighbors.

On *Figure 11* are shown the three densities for the eigenvector centralities (*A*: random; *B*: small-world; *C*: preferential attachment). The most different one, as expected, is the preferential attachment network: most nodes have little connections, while some nodes have really high values. The random and small world networks follow distributions close to normality.

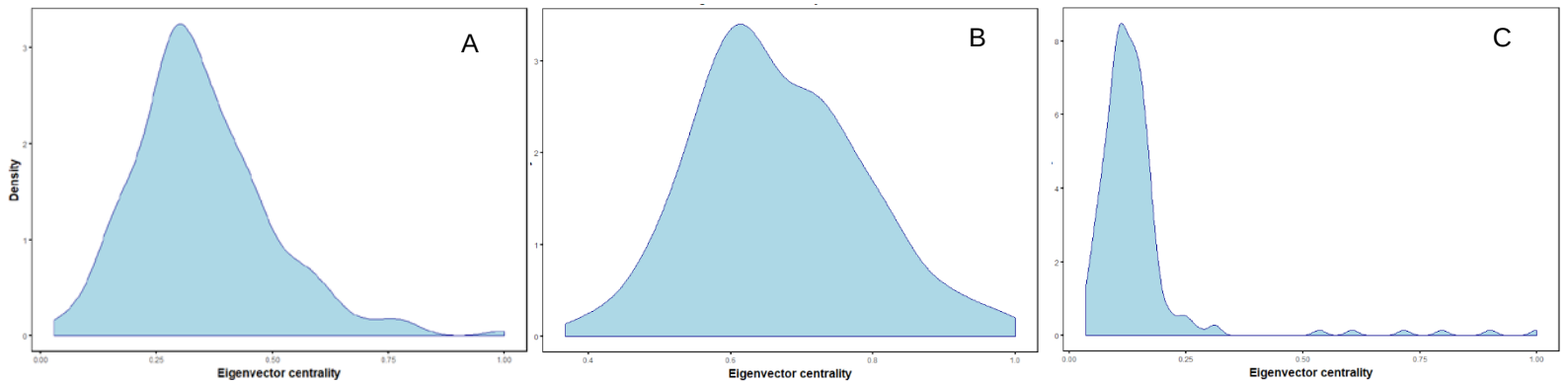

Figure 11

Rumor extintion

Finally, *Figure 12* shows the number of generations it takes for the rumor to extinguish for the three graphs, with 20 repetitions each. The small world network has the highest mean (close to 8, and is also the most uniform, while the other two have a smaller mean a little above 6 and both resemble less than the small world network to a normal distribution. This can be because 20 repetitions are too few, and I should have done more.

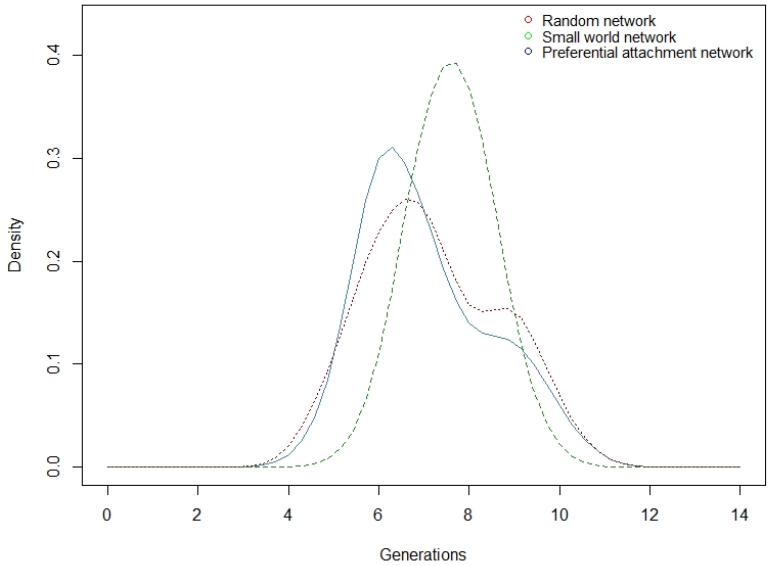

Figure 12