

SOCIAL NETWORK ANALYSIS

GROUP PROJECT ON

“SOCIAL SIMULATION”



Submitted By (GROUP 3) –

CHAHINE Carl
CAZILHAC Paul
BHURJI Prineet kaur
BARKETALLAH Salma
TIRUMALE LAKSHMANA RAO Kiran

Topic Description –

Social simulation is the application of computer-based methods and technologies to replicate human social behaviour in various environments and scenarios.

With the increasing popularity of social networks, it is becoming more and more crucial for the decision makers to analyse and understand the evolution of these networks to identify for example, potential business opportunities. Unfortunately, understanding social networks, which are typically complex and dynamic, is not an easy task. We thus need an effective and practical approach for simulating social networks.

Commonly, simulations are used to yield predictive data about what might happen in a real-world situation. Social simulations are also used to develop and test theories and to understand the implications of human behaviour.

Data Overview –

The project we selected for our Social Network Analysis class is about the diffusion of a friendship culture for a given network. To do this, we used two different datasets that we combined.

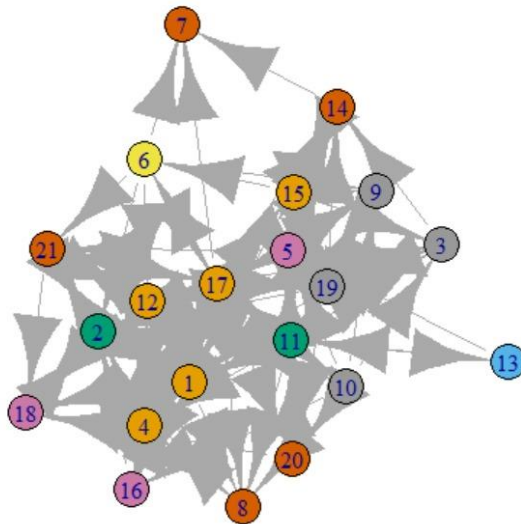
The First Dataset: 'Krack-High-Tec-edgelist-Friendship.txt'

This data is collected from the executives of a high-tech company. This company, which manufactures high-tech equipment on the West Coast of the United States, has just over 100 employees and 21 managers.

This data set is composed of three columns. The first two represent the nodes. Each manager is represented by a number in the 2 first columns Ego and Alter. The third column shows whether the managers are friends. It is a categorical encoded variable with 0 meaning no friendship and 1 friendship relation. As you can see in this summary table, both 'alter' and 'ego' columns have the same statistical values which is logical.

ego	alter	friendship_tie
Min. : 1	Min. : 1	Min. : 0.0000
1st Qu.: 6	1st Qu.: 6	1st Qu.: 0.0000
Median :11	Median :11	Median : 0.0000
Mean :11	Mean :11	Mean : 0.2313
3rd Qu.:16	3rd Qu.:16	3rd Qu.: 0.0000
Max. :21	Max. :21	Max. : 1.0000

The following graph shows how the managers are connected:



The Second Dataset: 'Krack-High-Tec-Attributes.csv'

It gives information about the managers in question meaning their age, their years spent in the company, their level, and their department.

When investigating the data, the following information has been found:

The average age of the managers is 40 years old; the youngest manager is 27 and the oldest is 62. The average seniority is more than 12 years and some managers have been in the company for only few months while others have been there for 30 years. For the department variable, the departments are encoded from 0 to 4. The managers are spread in all the departments as shown in this table:

Department	0	1	2	3	4
Number of managers	1	5	8	3	4

There are also 3 different levels, only 1 manager is in the first one, 4 in the second one and the next 16 in the third one. We can assume that the 1st level is the general direction and that it grows in a pyramidal form with 16 junior managers.

For further details on these characteristics here is the summary table for these attributes:

AGE		TENURE		LEVEL		DEPT	
Min.	:27.00	Min.	: 0.250	Min.	:1.000	Min.	:0.00
1st Qu.	:33.00	1st Qu.	: 7.500	1st Qu.	:3.000	1st Qu.	:1.00
Median	:37.00	Median	: 9.333	Median	:3.000	Median	:2.00
Mean	:39.71	Mean	:11.746	Mean	:2.714	Mean	:2.19
3rd Qu.	:43.00	3rd Qu.	:12.500	3rd Qu.	:3.000	3rd Qu.	:3.00
Max.	:62.00	Max.	:30.000	Max.	:3.000	Max.	:4.00

Project Approach –

In the code, the first part was used to read in the data of friendship and convert this data to a graph object. Then we used 2 methods to identify how quickly the culture of friendship would disperse in our dataset –

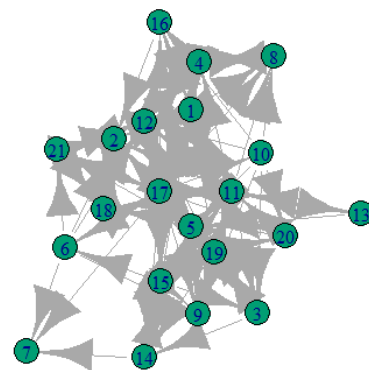
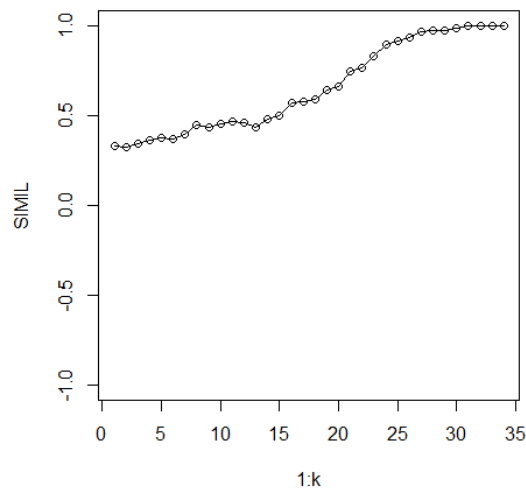
- 1) The first method involved a random utilization of 10 features across the individuals with a minimal level of one and a maximal level of five.
- 2) The second method involved a deeper split in the features. The first feature was related to the age and degree (i.e., people with a similar age or degree would relate quicker than others) and the second feature is related to the attitude (i.e., any liaison between the tenure, department or level would result in a quicker relation between the individuals). We were able to get this data thanks to the attribute file that we also read in in the beginning.

In this split we created 4 levels for the attitude feature which are a split by the following quantiles:

0.3, 0.6, 0.9 and 1 which means that any individual having a value of tenure under the 30th quantile of the total would have a discrete value of 0, while someone with a tenure that is between the 30th and 60th quantile would have a value of 1 and so on and so forth for all the variables. We also split the age and degree in a similar way (same quantile repartition).

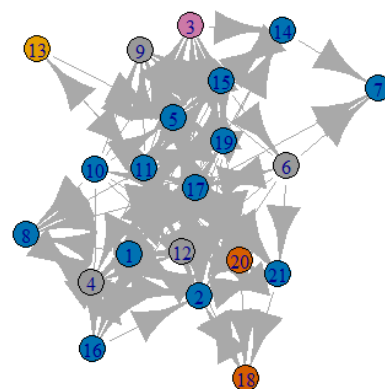
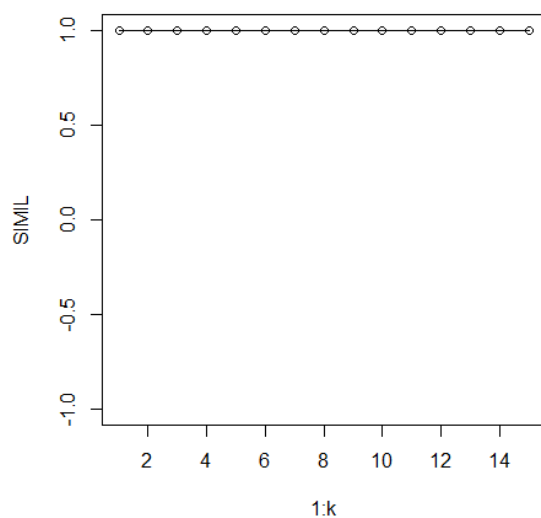
Next, we will identify how quickly the friendship culture disseminated in each case and what are the reasons for the convergence or divergence.

When using the method involving the random utilization of features across the individuals, below is the results we obtained:



As clearly shown by the above graph, after 32 iterations, we notice that all the individuals have the same colour in the graph, which means that they all converged to the same friendship culture.

Now if we look at the second method, ie the repartitions based on the quantiles, we get the following results:



Conclusion –

This shows that since the beginning, people having a similar repartition in the quantiles had their friendship culture spread except for 2 individuals: '**Individual 3**' and '**Individual 13**', and that is because they do not have similarities in any of the 2 features selected (age, degree & tenure, department and level).

Finally, in a work environment we would probably have a result similar to the final graph shown above, which shows that not all employees are going to blend and be friends no matter how long they stay together, while other would blend instantly.

REFERENCES –

- [Social simulation - Wikipedia](#)
- [Network simulation - Wikipedia](#)
- [What is social simulation? - Definition from WhatIs.com \(techtarget.com\)](#)
- [Social network simulator | Peter Cook \(peterrcook.com\)](#)
- [A simulation method for social networks | SpringerLink](#)
- https://ethz.ch/content/dam/ethz/special-interest/gess/social-networks-dam/documents/jss_netsim.pdf
- [Solve Social Phenomena with Simulations! \(digest ver\) - Bing video](#)
- [A Social Network Simulation In The Tidyverse | STATWORX](#)
- [A Social Network Simulation In The Tidyverse | R-bloggers](#)