The Class Network

Name: Paulo Junio Reis Rodrigues - RA: 265674

25/11/2020

1 Introdution

We asked the 15 students in the class (MO412) to tell us their hobbies. A total of 14 distinct hobbies were mentioned, and led to the network given in the accompanying file *class-network.tsv*. This is a bipartite network, because each link has a student in one extremity and a hobbie in the other. Both students and hobbies were anonymized, so students are identified by codes S1 to S15, whereas hobbies are identified by codes H1 to H14.

2 Algorithm

All programs were implemented using Python 3.8.5, with the Numpy 1.19.2, NetworkX 2.5, Matplotlib 3.3.2 and Pandas 1.1.4.

3 How to execute the script

In this project, we have one script (classNetwork.py), to run this script, you need to run this command:

```
python classNetwork.py
```

When the script finishes executing, it will generate the information / images from the bipartite network, and the two projections analysis. However, this script needs the file (class-network.tsv) to run.

4 Analyse

This class network is a bipartite graph, where students from one group and hobbies from the other, considering that each student may have more than one hobbie and students may also have hobbies in common. With this network we can define 3 views: First is a bipartite graph with 29 nodes and 55 links, second is a projection of students with 15 nodes and 83 links, and lastly, we have the projection of hobbies with 14 nodes and 50 links.

4.1 Bipartite graph

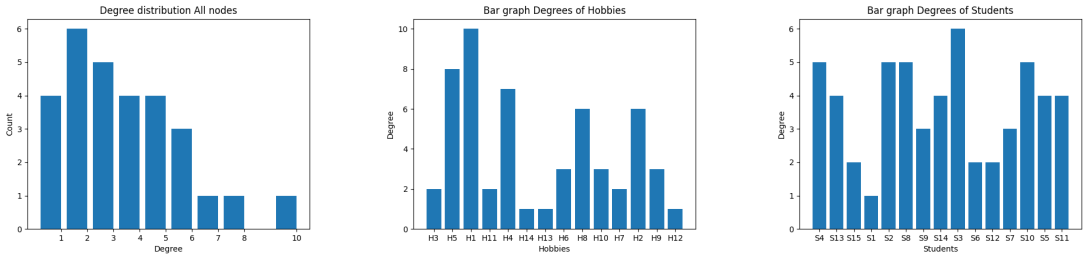
4.1.1 Degree distribution

To analyze the distribution of degrees in this graph, we need to visualize 3 parts, first, it is all nodes in the bipartite graph, the second is the nodes hobbies part and finally the nodes students part, the distributions are in figure 1. We can see, the most popular hobby is **H1** and only 1 student (S1) has one hobbie in the class.

4.1.2 Average degree

The average degree of this network, can be calculated by this formula:

$$\langle k \rangle = \frac{2L}{N}$$



(a) (b) (c)

Figure 1

Considering that L is the number of links and N is the number of nodes, for our case we will have this formula:

$$\langle k \rangle = \frac{2 * 55}{29} = \frac{110}{29} \sim 3.79$$

4.1.3 Average degree for a part

In this case, the average degree of part of the bipartite graph was calculated, this means that the average of hobbies done by students in the classroom is 3,928 and the average of students doing hobbies is 3,666.

4.1.4 Connected Components

The connected components are related to the number of subgraphs on the network. To calculate the components the command used *connected_components_number* from the **NetworkX** library. In this case, we find only one subgraph within the bipartite graph, that means that all students have at least 1 hobbie in common, therefore, there is no student who has only exclusives hobbies. This giant component has 29 nodes.

4.1.5 Average distance

Now we are going to calculate the average distance in this network, using the command *average_shortest_path_length* from **NetworkX** library to get the average distance, and the result is 2.699.

4.1.6 Clustering coefficient

The clustering coefficient captures the degree to which the neighbors of a given node are linked. Using the command *algorithms.cluster.clustering* from **NetworkX** to calculate all the clustering coefficients from every node in the graph. In this case, the neighbors have no links between them, because it is a bipartite network. Therefore, the clustering coefficient is 0 at all nodes in the bipartite graph.

4.1.7 Average clustering coefficient

In the previous item, we calculated the cluster coefficient for each node, now we will use the *algorithms.cluster.average_clustering* from **NetworkX** library to calculate the average clustering coefficient. The result is 0.

4.1.8 All results

In figure 2, will show all the results found in the analysis of the bipartite graph.

```
Name: All nodes
Type: Graph
Number of nodes: 29
Number of edges: 55
Average degree:   3.7931
Number of components: 1
Size of the giant component in graph 29
Average distance of graph 2.6995073891625614
Average Clustering coefficient 0.0
Average degree of Hobbies is 3.9285714285714284
Average degree of Students is 3.6666666666666665
```

Figure 2: All results for bipartite network

4.2 Hobbies projection

Can we create the hobbies projection, using the command *algorithms.bipartite.projected_graph* from the **NetworkX**, passing the all hobbies nodes.

4.2.1 Degree distribution

Considering the hobbies projection, we have the following distribution in figure 3:

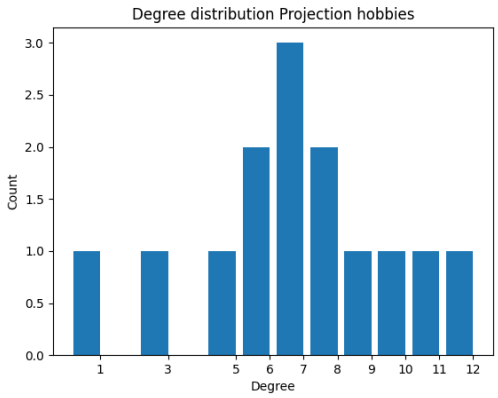


Figure 3: Hobbies projection degree distribution

4.2.2 Average degree

Considering the same formula as the average degree in the previous section, we can conclude that:

$$\langle k \rangle = \frac{2 * 50}{14} = \frac{100}{14} \sim 7.142$$

4.2.3 Connected Components

Considering the same command (*connected_components_number*) from the previous section, we can find only 1 component, with the size of 14.

4.2.4 Average distance

Using the same command (*average_shortest_path_length*) from the previous section, we find the result is 1.505.

4.2.5 Clustering coefficient and average clustering coefficient

Using the same commands (*algorithms.cluster.clustering* and *algorithms.cluster.average_clustering*), and get the this results, the clustering coefficient of all nodes it is on Figure 4, and the average clustering coefficient is 0.719. It is noted that **H14** has a clustering coefficient of 0, because as he is a hobbie of a single person, he therefore links only to 1 other hobbie.

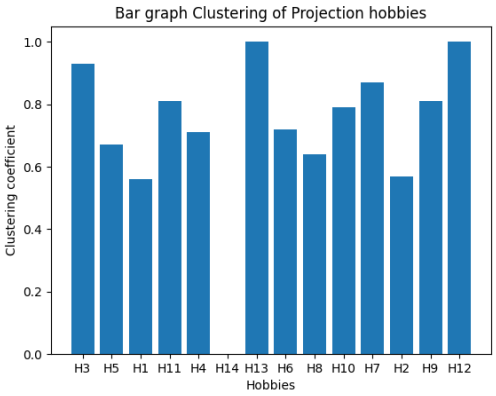


Figure 4: Hobbies projection clustering coefficient

4.2.6 All results

In Figure 5, will show all the results found in the analysis of the hobbies projection:

```
Name: Projection hobbies
Type: Graph
Number of nodes: 14
Number of edges: 50
Average degree:   7.1429
Number of components: 1
Size of the giant component in graph 14
Average distance of graph 1.5054945054945055
Average Clustering coefficient 0.7197381983096268
```

Figure 5: All results for hobbies projection

4.3 Students projection

Can we create the students projection, using the command *algorithms.bipartite.projected_graph* from the **NetworkX**, passing the all students nodes.

4.3.1 Degree distribution

Considering the students projection, we have the following distribution in figure 6:

4.3.2 Average degree

Considering the same formula as the average degree in hobbies projection section, we can conclude that:

$$\langle k \rangle = \frac{2 * 83}{15} = \frac{166}{15} \sim 11.066$$

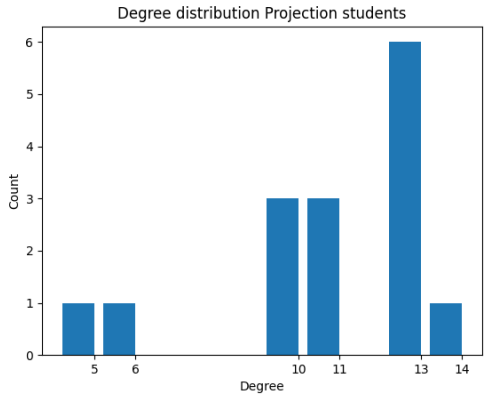


Figure 6: Students projection degree distribution

4.3.3 Connected Components

Considering the same command (*connected_components_number*) from the hobbies projection section, we can find only 1 component, with the size of 15.

4.3.4 Average distance

Using the same command (*average_shortest_path_length*) from the previous section, we find the result is 1.209.

4.3.5 Clustering coefficient and average clustering coefficient

Using the same commands (*algorithms.cluster.clustering* and *algorithms.cluster.average_clustering*), and get the this results, the clustering coefficient of all nodes it is on Figure 7, and the average clustering coefficient is 0.889. All clustering coefficients are very high, showing that we students have several similar profiles, with similar hobbies style.

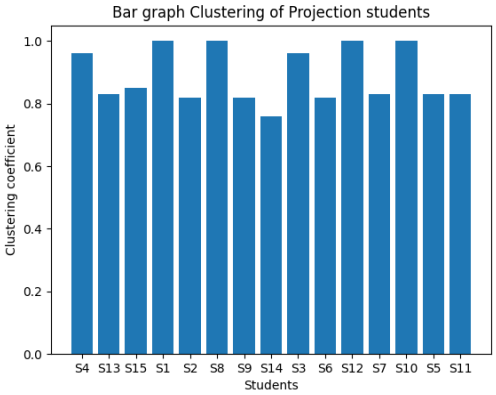


Figure 7: Students projection clustering coefficient

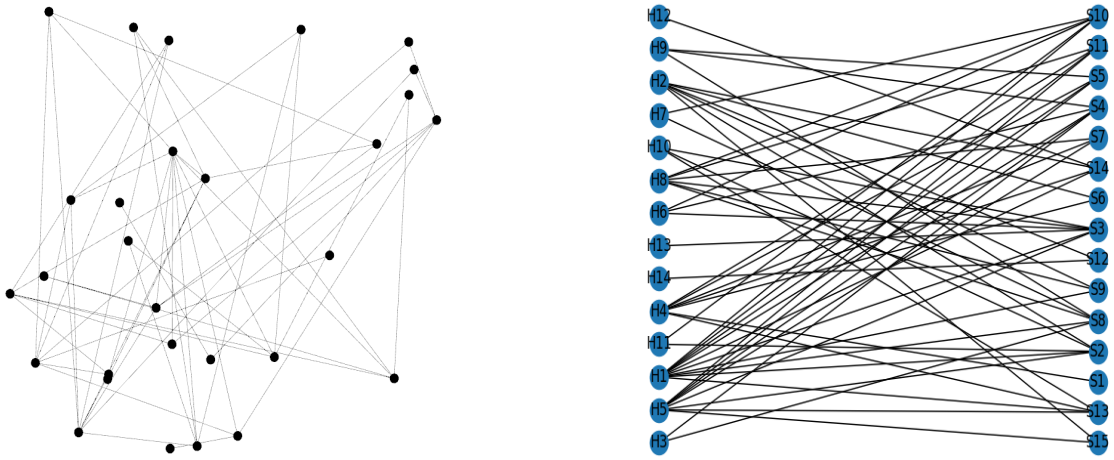
4.3.6 All results

In Figure 8, will show all the results found in the analysis of the students projection:

5 Plots

```
Name: Projection students
Type: Graph
Number of nodes: 15
Number of edges: 83
Average degree:  11.0667
Number of components: 1
Size of the giant component in graph 15
Average distance of graph 1.2095238095238094
Average Clustering coefficient 0.8891094091094093
```

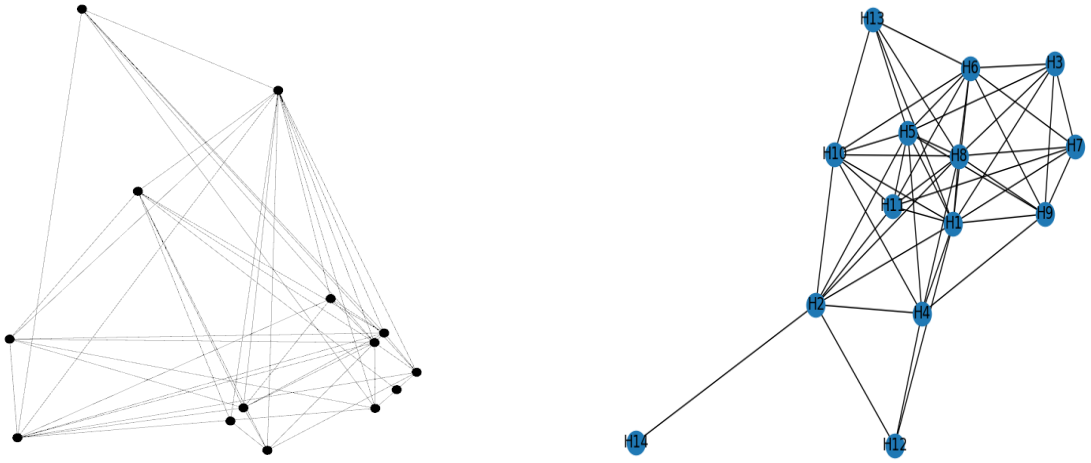
Figure 8: All results for students projection



(a) Gephi Plot

(b) NetworkX Plot

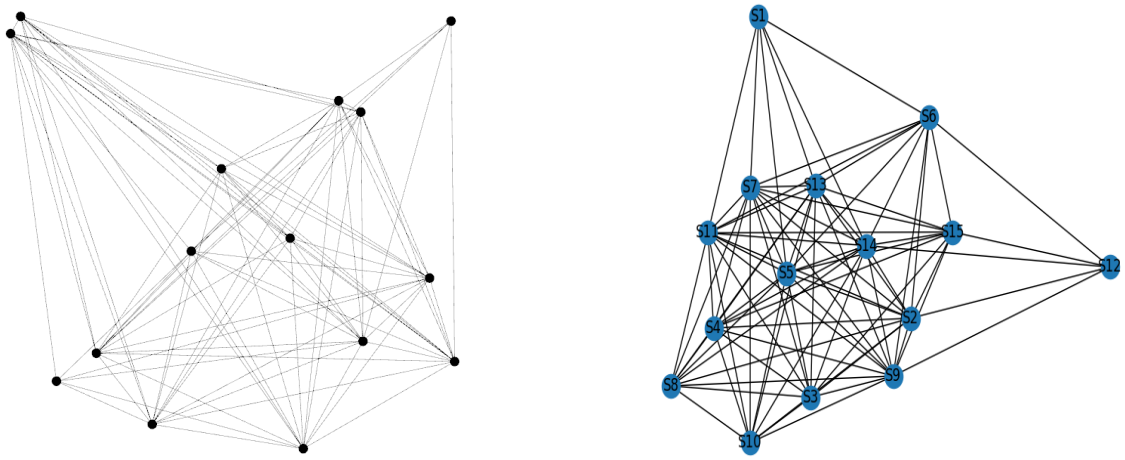
Figure 9: Bipartite graph plots.



(a) Gephi Plot

(b) NetworkX Plot

Figure 10: Hobbies projection plots.



(a) Gephi Plot

(b) NetworkX Plot

Figure 11: Students projection plots.