Complex social network Analysis: Case of interaction between student at the restaurant

Cedric Lontsi SAADIO (cedric.l.saadio@aims-senegal.org)
African Institute for Mathematical Sciences (AIMS)

Supervised by: Prof. Franck kalala Mutombo
Institute of Supervisor, Country

March 25, 2020

Assignment in Complex Networks

Abstract

Coupled biological and chemical systems, neural networks, social interacting species, the Internet and the World Wide Web, are only a few examples of systems composed by a large number of highly interconnected units.The first approach to capture the global properties of such systems is to model them as graphs whose nodes represent the units, and whose links stand for the interactions between them. In this document, we analyse the social network of interaction between students of African Institute for Mathematical Sciences(AIMS) Senegal cohort 2019-2021 during their launch time. From the result of that analyse, we discover that the level of interaction between students is very weak. We got that when we pick randomly a student, there is 18.8 % of chance that his neighbour are knowing each other. We also discover that there are 8 communities in the network.

Declaration

I, the undersigned, hereby declare that the work contained in this research project is my original work, and that any work done by others or by myself previously has been acknowledged and referenced accordingly.

Scan your signature

Lontsi Saadio Cedric,March 25, 2020

Contents

1. Introduction

A network is a graph denoted by $G(V, E, \gamma)$ where V is the set of nodes, E is the set of link (edges) between a pair of node $E = \{(u, v) \in V \times V / u \in V, v \in V\}$ and γ is a function $\gamma : V \times V \to \mathbf{R}_+$, sometime, γ is seen has the function of weight attribution. A network become complex when he had a large number of node and edges and it's not easy to visualize and analyse. Complex network is nowadays widely apply in many domain such as data science(network science),web (interconnection between server in the web), transport(transportation network) , medicine (brain network analysis), chemistry (interaction between atom),sociology,etc. In the domain of sociology, complex network can help to analyse interaction between individuals in the society such as sexual relationship, interaction between people from the same organisation, community religious,etc.

African Institute for Mathematical Sciences(AIMS) in Senegal is a panafrican institute of mathematical sciences which select each year excellent student in many countries across Africa to pursued an excellent training in mathematics in meet to become the next Einstein . At launch time, those students used to sit together in group on different table. When a student take his food, he rigorously choose a table where to sit, the choice of the table because is based on a preference on those who are sitting on that table. Hence Each student have his own colleague with whom he would like to discuss while eating during the launch time.

Let represent each student by a node and let define the interaction(discussion) or preference as a link between a pair of students, in that case, we come out with a complex social network which represent the interaction between student during their launch time. By having that network, we can ask ourselves, which information can we extract from that network by analysing it?

In this work, we are going to analyse the social network of interaction between AIMS student while there are taking their launch.

Our work is divided as follow, in the first part of our work, we will present the necessaries tools that we used and the topology of our network, in the second part, we are going to do an analysis of the network.

2. Tools and Topology of the network

2.1 Necessaries Tools

There exist many tools to implement and visualize network such as **networkx**(python libraries); **igraph**(Python, C, R); **pajet**, (for vizualisation); **Gephi**(for vizualisation). In this work, we will use pyhton as the programming language with his libraries such as networkx, matplotlib(with the method pylab or pyplot), Numpy, Collection(with the method Counter).

2.2 Topology of the network

In this section, we present the Topology of the network. By using networkx, the following commands help us to visualize the topology o the graph: Here is the output of the command:

```
options = {
          'node_size': 100,
          'width': 3,
          }
G = nx.Graph()
G.add_edges_from(List_of_nodes)
plt.figure(figsize=(15,15))
nx.draw(G,with_labels=True,**options)
plt.show()
```

Figure 2.1: Code for create the network using networkx

The network is composed of **58 nodes** and **93 edges**. The 58 nodes represent the number of student at AIMS Senegal cohort 2019-2021 and 93 edges represent the number of interactions between those students.

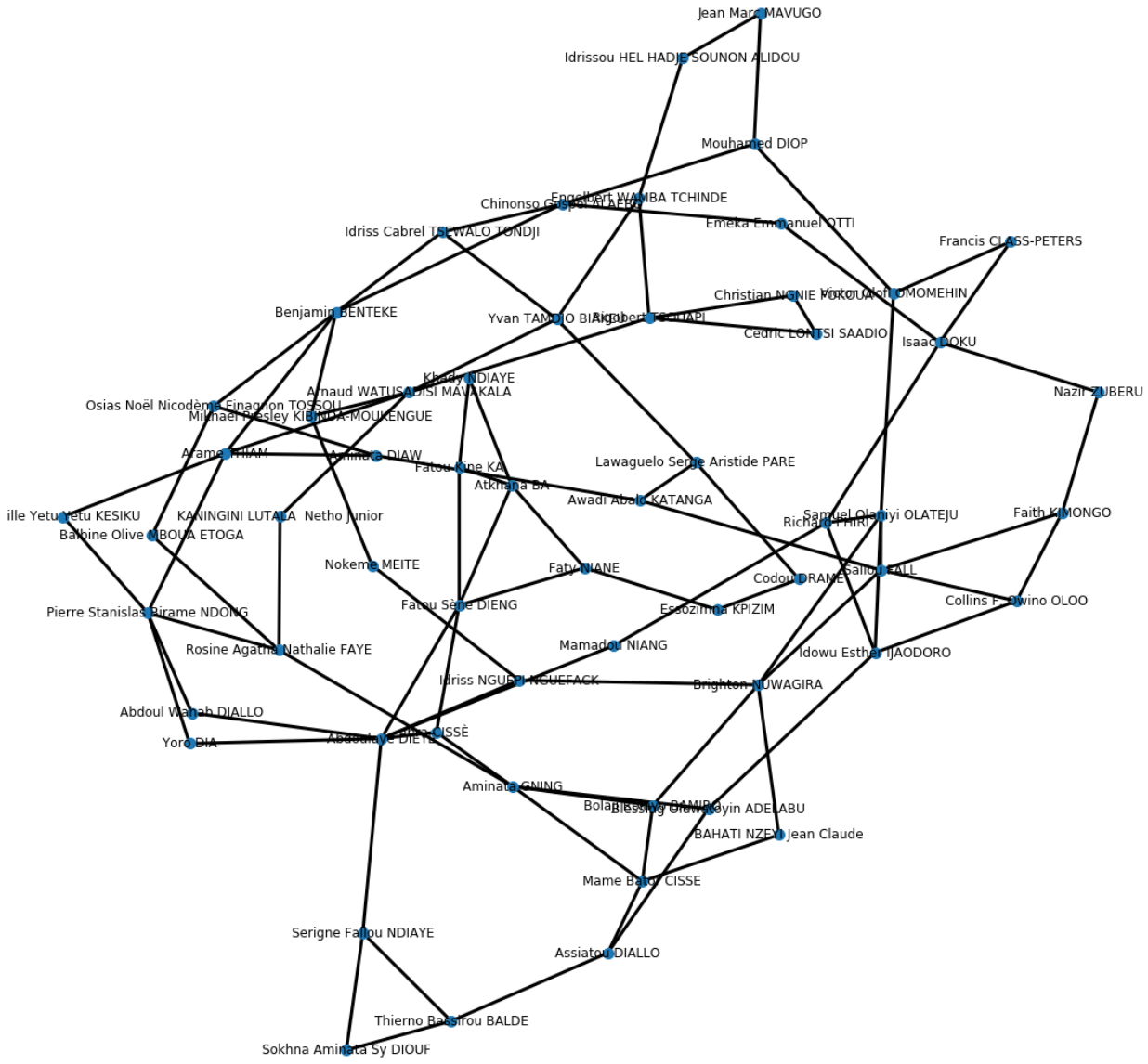

Figure 2.2: Social network of the interaction between students in launch time At AIMS Senegal.

3. Network analysis

3.1 Degree distribution

The degree of a node V is the number of direct link between him and his neighbours. In our network, The node named **Abdoulaye DIEYE**, is the node which have the maximum degree(7), it means he is the student with who most of student would like to sit width while eating their meal.So he is the most important student in the network.

By looking at the degree description (in the table below), we can see that the mean degree of the network is 3.20, this means most of students have tendency to interact with 3 students while they are eating.The std of the network is 1.16, so there is not a large dispersion between the number of interaction between students. The minimum degree is 2, so each student is connected to at least 2 students(fig 2.2).

Mean degree	std	Min degree	Max degree	Quartile 1(25%)	Quartile 2 (50%)	Quartile 3(75%)
3.20	1.16	2	7	2	3	4

Table 3.1: Description table of the degree of the restaurant network

We compute The degree distribution of the network to see the way student interact between themselves. Here(Figure 2.3) is the source code for create the plot of degree distribution. The output is (Figure 2.4)

```
Nk= Counter(resultdegree.degree)
K= sorted(list(Nk))
PK = [Nk[i] for i in K]
plt.bar(K,PK,color="#fd722c")
plt.plot(K,PK,color='green')
plt.title("The degree distribution")
plt.xlabel("degree k")
plt.ylabel("Count")
plt.show()
```

Figure 3.1: Degree distribution source code of the network of interaction between Aims student in launch time

By looking at the plot, we can say that the degree distribution of our network is heterogeneous,hence the degree follow a power law distribution. An interpretation of this result is that there is a large number of students who don't interact together.We can see on the plot that degree 2 and degree 3 have the high number of representation, this can signify that most of student like to sit with two or three colleagues while there are eating.

3.2 Clustering

The global clustering of the network is 0.1881 in other word, if we take randomly one student, the is 18.81% of chance that those who are eating with him interact each others. This show again how level the relationship between students is weak.

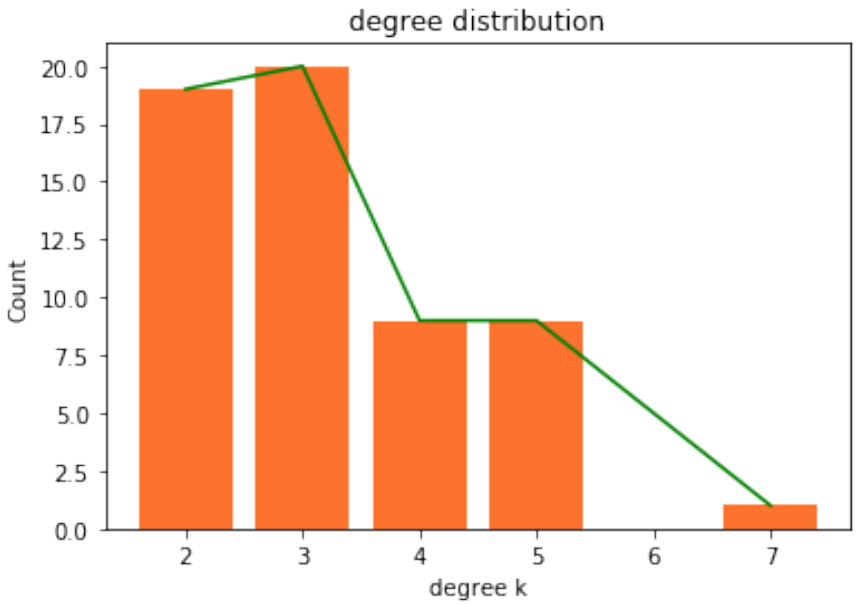

Figure 3.2: Degree distribution of the network of interaction between Aims student in launch time

3.3 Diameter,radius and center of the network

The diameter of the network define the largest distance between two students.In networkx we use the command **nx.diameter(G)**. The result is **9**.An interpretation is that when we pick randomly two students, the maximal number of link one go through when one move from student A to to student B is 9.

The radius of the network is the smallest distance between two students, the command to use in networkx is **nx.center(G)** and the result is 5.

The center of the graph is the network is **Nokeme meite** because she is the student who have the lowest eccentricity which is equal to the radius.So Nokeme is the most important and influence student in our network.It means if there is a spreading information, she is best person by whom the information can easily spread.

3.4 Community detection

In this section, we are going to detect communities that can be form in our network.We use two techniques for community detection which are community graph and hierarchical clustering.

In network science,a community is a group of nodes of a network in which nodes are connected to each other more densely than to the outside. By using the community graph through the command **networkx.algorithms.community**, we obtain **8** communities, it means there are 8 groups of students who like to sit together. We can visualize those communities on the figure below (figure 3.3). By looking the figure, we can say that student use to sit together according to their origin's country, language or culture. For example, The community which contain Nazir Zuberu is English speaker in majority(only saliou and mamadou Niang who are not English speaker). Those who are in the cluster of Faty Niane

are coming from the same country thus can share the same culture so they are more comfortable when there are sitting together. In this case, it's very easy to detect the member of the community, hence the analysis become easy.

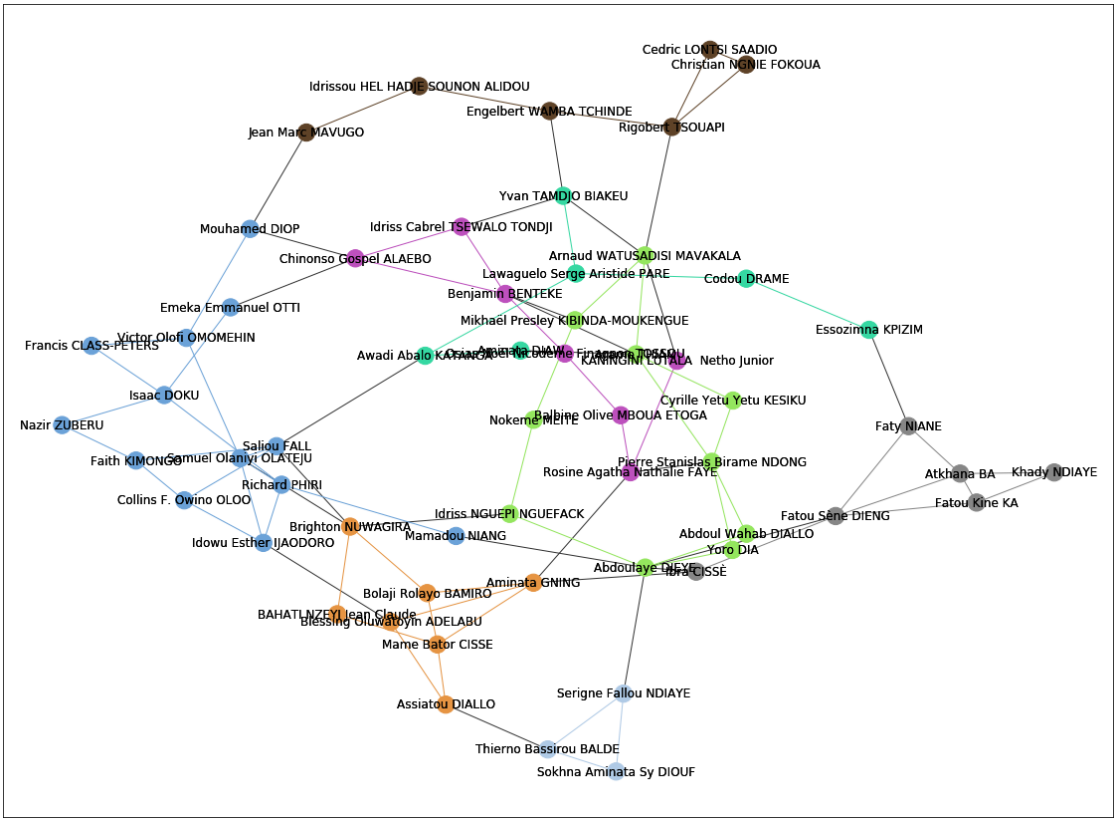

Figure 3.3: Visualization of communities in the network using community graph

If we use the hierarchical clustering method(with "single" property) which is a statistical data clustering technique that can be used to cluster pairs of nodes which are nearly structural equivalent into positions, we come out with **7** communities(fig3.4) but with less specificity. We don't have enough of information on those community in other word we can detect individuals from those communities.

3.5 Betweenness centrality

In this part, we analyse the ability of each person to be the way of communication between communities.

We compute those values with networkx by using the command **nx.betweenness_centrality(G)** and we present in the form a dataframe using pandas. In the figure 3.5, we can see the top 5 of the descending sorted list of student and we remark that the student **Abdoulaye DIEYE** is on top (he has the highest degree) with **0.228718** as the value of betweeness centrality. So Abdoulaye is the student who interconnected most of communities after come the student **Arnaud WATUSADISI MAVAKALA** with the value **0.142153**.

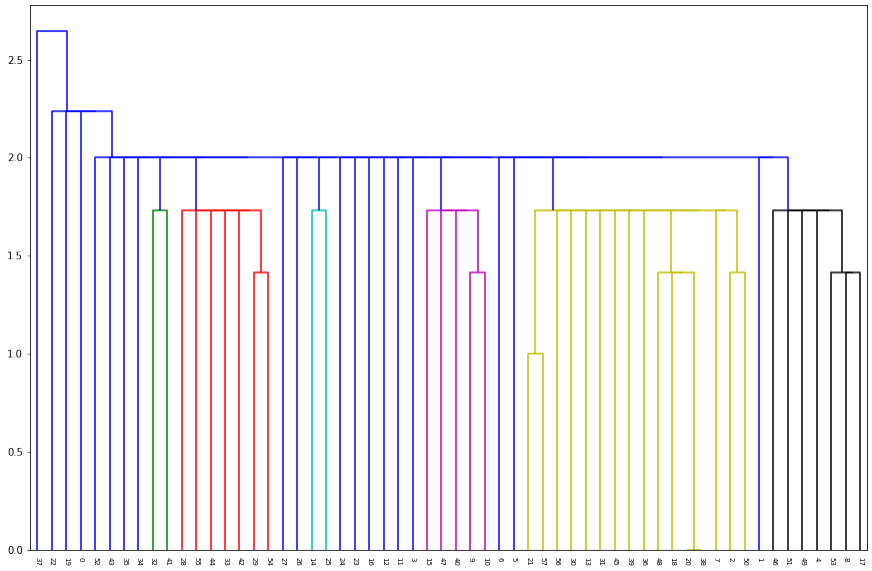

Figure 3.4: Visualization of communities in the network using hierarchical clustering

	Name	betweeness
37	Abdoulaye DIEYE	0.228718
3	Arnaud WATUSADISI MAVAKALA	0.142153
35	Aminata GNING	0.130493
28	Fatou Sène DIENG	0.130184
5	Arame THIAM	0.115752

Figure 3.5: Betweeness centrality of the graph

3.6 Closeness centrality

In this section, we are looking for the student who is more close to communities and from whom we can easily access a community. We compute those values with networkx by using the command **nx.closeness_centrality(G)** and we present in the form a dataframe using pandas.In the figure 3.6, we can see the top 5 of the descending sorted list of student and we remark that the student **Brighton NUWAGIRA** is the top of the list with the value **0.303191**. It means he is the most closest student from community and from him, a community can easily assess other communities. After him, the next is **Idriss NGUEPI NGUEFACK**.

:

	Name	betweeness
22	Brighton NUWAGIRA	0.303191
36	Idriss NGUEPI NGUEFACK	0.300000
37	Abdoulaye DIEYE	0.300000
35	Aminata GNING	0.295337
5	Arame THIAM	0.293814

Figure 3.6: closeness centrality of the graph

Conclusion

In summary, our work was to implement and analyse the network of interaction between AIMS Senegal student at launch time. We first of all look at the topology of the network,after we start the analyse by describe the degree (mean degree, std, min,max,...) where we saw that most of students are directly connected to 3 others students, we continue the analyse by looking at the degree distribution where we discover that the degree of the network is heterogeneously distributed. When we compute the global clustering, we discover that if we pick randomly one student in the graph, his two neighbours will know each other. We also compute the diameter, the center and the radius of the network and we discover that, the maximal distance between two students is 9(the diameter), the most important individual in the graph is Nokeme meite (center).We detect 8 communities of students and most of those communities where form based on language, culture and origin's country. We end the analysis with the computation of the betwenness and the closeness centrality node. In this work, we only focus on the interaction between students, how can be the network if we include the staff and tutors?

References

Igraph. Igraph documentation http://igraph.org/redirect.html.

Networkx. Networkx documentation https://networkx.github.io/documentation/stable/.

Gephi. pajek documentation https://gephi.org/.

Pajek. pajek documentation http://vlado.fmf.uni-lj.si/pub/networks/pajek/.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. Complex networks: Structure and dynamics. *Physics Reports*, 424(4):175 – 308, 2006. ISSN 0370-1573. URL http://www.sciencedirect.com/science/article/pii/S037015730500462X.

Estrada, E. *The structure of complex network, theory and application*. Oxford, 2011.

Mohammed Zuhair Al-Taie, S. K. *Python for graph and network analysis*. Springer, 2017.

Newmann, M. *Network, an introduction*. OXFORD, 2010.