# *Terror Affiliation*

## *Final Project Report*

*Team: Johanna Fransson, Jared Godinez, Giang Duong*

*https://github.com/johannafransn/Social-Network-Analysis-Terror-Affiliation*

# *Introduction*

Social network analysis can provide valuable information, especially when analyzing a network of specific groups of people. You can learn how the group communicates and how they are related with one another. For example, networks of terrorist groups, such as the groups involved in the September 11th terrorist attacks and other attacks, can be analyzed and can thus be used to determine how the individuals had known each other, how they coordinated their plots, and how they were ultimately able to carry out their plans. For our project, we used different approximations and calculations to solve the questions of who were the most important individuals in the network, and to discover the centralities of the networks. In order to implement these networks with visuals, we will be using Jupyter Notebook, Python, and libraries such as networkx, pandas, matplotlib, and seaborn.

# *Objective*

Our objective for our project is to analyze a dataset that we had found that relates to 9/11 terrorist groups.The dataset that we used for our project can be found at: http://www.casos.cs.cmu.edu/tools/datasets/external/index.php. This data provides data on the people who were directly involved in the terrorist attacks, and also people who were indirectly involved. In order to allow a much more efficient and easier usage of the data, we will be first extracting specific information from this dataset, and then extracting that information into a CSV file. Once we have created the CSV file, we utilize the libraries networkx, pandas, matplotlib, and seaborn.

# *Data Gathered*

*Dataset in XML before clean-up:*

```
<link source="Ahmed Alnami" target="Saeed Alghamdi*"/>
<link source="Ahmed Alnami" target="Hamza Alghamdi"/>
<link source="Ahmed Alghamdi" target="Hamza Alghamdi"/>
<link source="Hamza Alghamdi" target="Ahmed Alnami"/>
<link source="Hamza Alghamdi" target="Ahmed Alghamdi"/>
<link source="Hamza Alghamdi" target="Saeed Alghamdi*"/>
<link source="Hamza Alghamdi" target="Ahmed Al Haznawi"/>
<link source="Hamza Alghamdi" target="Mohand Alshehri*"/>
<link source="Ahmed Al Haznawi" target="Saeed Alghamdi*"/>
<link source="Ahmed Al Haznawi" target="Hamza Alghamdi"/>
<link source="Ahmed Al Haznawi" target="Ziad Jarrah"/>
<link source="Mohand Alshehri*" target="Hamza Alghamdi"/>
<link source="Mohand Alshehri*" target="Fayez Ahmed"/>
<link source="Fayez Ahmed" target="Mohand Alshehri*"/>
<link source="Fayez Ahmed" target="Marwan Al-Shehhi"/>
<link source="Ziad Jarrah" target="Ahmed Al Haznawi"/>
<link source="Ziad Jarrah" target="Marwan Al-Shehhi"/>
<link source="Ziad Jarrah" target="Mohamed Atta"/>
<link source="Marwan Al-Shehhi" target="Fayez Ahmed"/>
<link source="Marwan Al-Shehhi" target="Ziad Jarrah"/>
<link source="Marwan Al-Shehhi" target="Mohamed Atta"/>
<link source="Marwan Al-Shehhi" target="Abdul Aziz Al-Omari*"/>
```

*Dataset in csv format after clean-up:*

```
From,To
Majed Moqed ,Hani Hanjour
Khalid Al-Mihdhar ,Hani Hanjour
Khalid Al-Mihdhar ,Nawaf Alhazmi
Hani Hanjour ,Majed Moqed
Hani Hanjour ,Khalid Al-Mihdhar
Hani Hanjour ,Nawaf Alhazmi
Nawaf Alhazmi ,Khalid Al-Mihdhar
Nawaf Alhazmi ,Hani Hanjour
Nawaf Alhazmi ,Salem Alhazmi
Nawaf Alhazmi ,Ahmed Alnami
Nawaf Alhazmi ,Saeed Alghamdi
Nawaf Alhazmi ,Hamza Alghamdi
Salem Alhazmi ,Nawaf Alhazmi
Ahmed Alnami ,Saeed Alghamdi
Ahmed Alnami ,Hamza Alghamdi
Ahmed Alghamdi ,Hamza Alghamdi
```

## Import the dataset and display general attributes

```
In [1]:
import networkx as nx
import matplotlib.pyplot as plt
%matplotlib inline
import pandas as pd
import seaborn as sns
import matplotlib.colors as mcolors


df = pd.read_csv("9_11_hijackers.csv") #load our dataset
```

```
In [2]:
df.head() #display the head of our dataset
```

| | From | To |
|---|---|---|
| 0 | Majed Moqed | Hani Hanjour |
| 1 | Khalid Al-Mihdhar | Hani Hanjour |
| 2 | Khalid Al-Mihdhar | Nawaf Alhazmi |
| 3 | Hani Hanjour | Majed Moqed |
| 4 | Hani Hanjour | Khalid Al-Mihdhar |

*Code to display information of the network:*

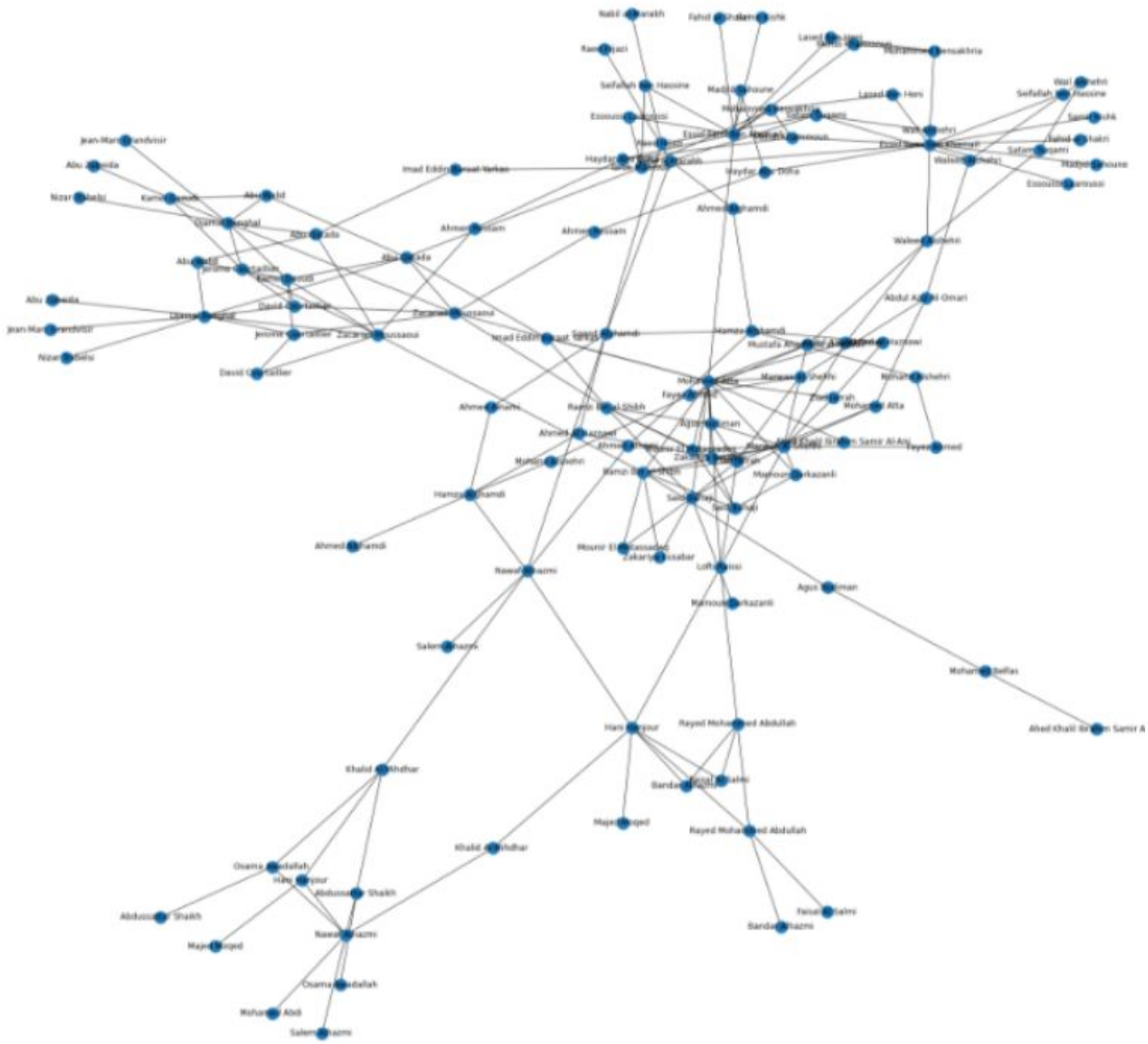
```
In [3]:
#read the edgelist
mygraph = nx.from_pandas_edgelist(df, source="From",target="To")

#display info about type, node/edge count avg degree etc
print(type(mygraph))
print(nx.info(mygraph))
```

```
<class 'networkx.classes.graph.Graph'>
Name:
Type: Graph
Number of nodes: 114
Number of edges: 187
Average degree:   3.2807
```

Figure A.
Source:
Software: Python 3.0

*Visual of the Network*

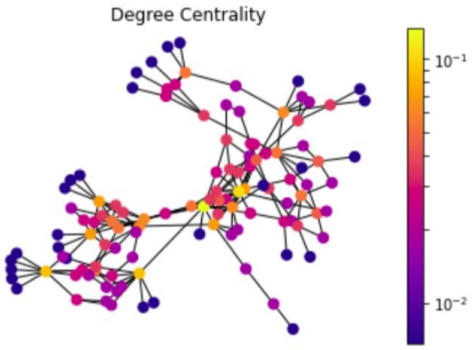


# *Observations*

Here we display our resulting network centralities. The three centralities we measured were the degree centrality, betweenness centrality, and eigenvector centrality.

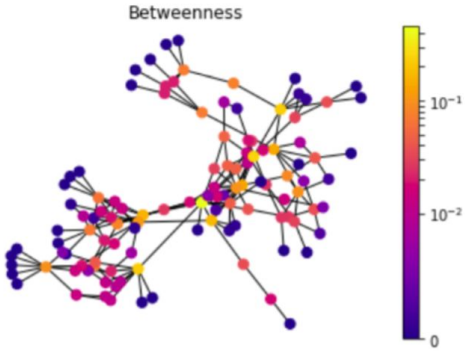
## *Degree Centrality*

```
[32]: pos = nx.spring_layout(mygraph, seed=675)
      draw(mygraph, pos, nx.degree_centrality(mygraph), 'Degree Centrality')
```



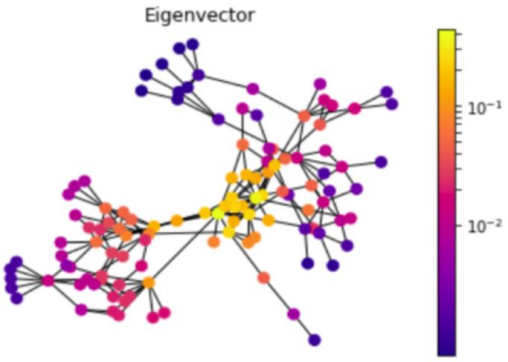

## *Betweenness Centrality*

```
pos = nx.spring_layout(mygraph, seed=675)
draw(mygraph, pos, nx.betweenness_centrality(mygraph), 'Betweenness')
```




## *Eigenvector Centrality*

```
pos = nx.spring_layout(mygraph, seed=675)
draw(mygraph, pos, nx.eigenvector_centrality(mygraph), 'Eigenvector')
```



# *Results*

One of the features that stood out the most for us was the most affluent node in the network. Based on our intuition, it seemed clear from just the appearance of the network, that Mohamed Atta would be the most critical node within the network. In the Eigenvector Centrality figure, we can see that the most yellow dot in the center of the network between the two major network components, is the influential node in the network, based on the measurement bar on the right side of the image. Along with being the most influential node, Mohamed appeared to have the most control over the network's information as well. This is due to the fact that Mohamed had the highest betweenness centrality, which can also be seen in the Betweenness Centrality image. What we were also able to point out was how sparse and spread out the network was in terms of the levels of the networks from Mohamed Atta to the outermost nodes. There were two major components in the network, and this could have been due to the fact that many of the members of the network held weak ties with each other, which allowed potential damages to the network to be minimal if someone in the network were to be captured or compromised. This discovery also serves as reinforcement to why Mohamed Atta was such a critical and influential node in the network, as he served as a bridge to many different people in the network.

# *Conclusion*

We can conclude from our results, that networks of terror groups such as the one we analyzed do not have many members form ties outside of their immediate ties, which the majority of the time was only one other tie in the whole network. When we had found out about this, we realized that it can be difficult to determine exactly whether a tie between nodes in the network were strong or weak. What was also interesting about the dataset was that it also included the individuals that were not directly involved in the attacks. This also allowed us to make observations and conclude which individuals were subject for investigation, and to not immediately claim guilt for anyone else that may possibly be linked to the group. This also allowed us to make the analysis of which specific people were most likely involved in the attacks, and which were not. The people in the outermost nodes of the network had mainly been those who were not involved directly in the attacks. As just as we expected, Mohamed Atta, the most influential and important in terms of communication and information between all nodes in the network, was one of the pilots that had hijacked the planes during these terror attacks.