Analyzing the Network of The ISIS Tweets in Twitter Using Social Network Analysis

Nupur Roy Chowdhury
14395901
nr572@drexel.edu

Abstract: Islamic State of Iraq and the Levant known as ISIS one of the terrorist Organizations which is a global threat controls Syria and most of Iraq and Israel. Analysis of this terrorist network is essential for discovering their propaganda. ISIS has a constant stream of propaganda for which it is using twitter 90% of the time to spread violence. ISIS consists of 70,000 twitter accounts globally out of which around 200,000 tweets are being posted on twitter every day. Studying the network of ISIS helps in developing counter strategies in order to blunt its outreach efforts effectively as it is using digital strategy to crowdsource. Using social network analysis and tools we will be able to explore the complex networks to recognize patterns which were previously unreadable. Through this we will be able to extract the hidden relationships, their strategies and the user interactions.

1.Introduction:

The ISIS is an organized terrorist organization that consists of collaboration between all the terrorist in that organization which forms the basis of the relationships which these terrorists share with each other. The ISIS is a force which takes into consideration almost 80% - 90% of the digital platform for its strategies. Where twitter is the cornerstone of this group's digital strategy.[1]

To investigate this network of the terrorist organization the following tasks are to be performed:[1]

Subgroup Detection: The different members of the terrorist organization perform different roles. It is important to identify the groups that perform these actions as this would help in identifying the important terrorists who leads this network. As the ISIS network consists of six different types of roles which are carried out by the user accounts in twitter who are known as actors to help identify the relationships between each type of the accounts.

Identification of important actors and their roles in the Network: As, we have seen that ISIS has six different types of roles which are played by the members of this network. Hence, it is important to identify the important actors of these roles who are playing the lead and connecting the network and building it together. The approach here will be to extract each user with their unique usernames which acts as a node. The node will be then influenced by the scale consisting of the combination of the number of followers and the number of tweets they tweet from their account. This combination helps
us to ensure that active and popular users has been identified. Currently this relation between each actors is yet to be fixed, as one relation criteria might be to scrape the tweets of users having maximum followers and then link the nodes via this metric with numerous mentions increasing the weight of an edge between the actors where they either have direct communication or by retweeting the tweets of the actors.

Discovery of the Interactions and Patterns between the actors: We need to find the pattern in which these actors communicate like using any specific kind of hashtags which the related actors uses or is their any common actor between each actors who gets tagged in the tweets and re-tweets and they collectively form a pattern through which interactions takes place. It is also based on qualitative methods of aggregated pairwise occurrence of the tweets for each actor, which can then be normalized by the sum of all tweets from all the actors that are being compared with each other. This will help in tweets which are exceedingly large and their actors and the actors with minimal tweets activity. To this the actor will be the main identifier. This will generate the undirected weighted graphs.

2. Terrorist Social Network: ISIS (Literature Review)

The ISIS is an organized terrorist organization that consists of collaboration between all the terrorist in that organization which forms the basis of the relationships which these terrorists share with each other. The ISIS is a force which takes into consideration almost 80% - 90% of the digital platform for its strategies. Where twitter is the cornerstone of this group's digital strategy.[1]

It consists around 46,000 - 70,000 twitter accounts all over the world promoting hate messages. An estimated follower of around 21,000 are English speaking language alone. There is around 2,000 over-performers accounts that tweet at a bursts of 50 or more tweets in a single day and each of these performers or actors again have an average of 1,004 followers or actors which results in around 200,000 tweets each day by all of these actors in spreading hate messages and contents. [1]

The messaging style of ISIS can be broken down into two types:[1]

- Content for Recruitment
- Content for its enemies

Content for Recruitment: This includes breaking news from the front headings and encouraging the youth to take matters into their own hands. Their online magazine called "Dabiq" describes "The Management of Savagery", video's such as the "Mujetweets" which function like mods for games which allows the players to live the life of an ISIS recruit going around shooting and blowing things up and their religious melodies called as nasheeds.[1]

Content for its enemies: This includes feature films threating countries of terror attack if they attempt to intervene in the regions occupied by them. Based on the twitter data there are six different types of roles which can be identified as: [1]

1. Reporters: users who convey the breaking news by ISIS.
2. Reconnectors: users whose job is to retweet the usernames of the violent extremists whose accounts were suspended before but are back with new names.
3. Intellectuals: who uses the philosophies, economics and political theory to justify their doings.
4. Fanboys: who have their profile picture as the ISIS and celebrates its victories.
5. Recruiters: who does the recruitment, they have private accounts where after a potential actor has been referred does direct messages to them and then takes it through encrypted channels.
6. Mujahideen: The actors posted as heroes as they are the ones who are fighting on the frontlines.

A social network is typically represented by a weighted graph *G* = (*V, E; w*), where *V* corresponds to the set of nodes, *E* is the set of links, *w* is a function mapping each link (u, v) € E to a weight w_{uv} in the range [0,1] that indicates the strength of association between *u* and *v*. Each node, *v*, is corresponding to a person, which is a terrorist in a terrorist social network (TSN). A link between two nodes (terrorists), *(u,v)*, represents that there are some kinds of relationships between the corresponding terrorists, *u* and *v*. The weight w_{uv} is determined by the number of types of relationships existing between *u* and *v*. Two terrorists can be related through different types of associations. We have heuristically assigned an importance score s_r to each type of relationship *r* and compute a total score s_{uv} for each link (*u*, *v*) as the total score of the relationships between *u* and *v*, i.e.,[20]

$$s_{uv} = \sum_{r \in R(u,v)} s_r$$

where *R(u,v)* denotes the set of relationships existing between *u* and *v*. The link weight w_{uv} is then computed as the normalized link score, i.e.,[20]

$$w_{uv} = \frac{s_{uv}}{\max\limits_{u,v \in V}(s_{uv})}$$

Based on the data set, the resulted terrorist social network consists of a total of 3328 nodes and 5331 links.

3. Visualizations of the Network found:

The computation of initial node coordinates and sizes are the most important steps in presenting the terrorists and their relationships, represented as a weighted graph *G* = (*V, E; w*), on a two-dimensional space. A mapping of each node v ∈V of the terrorist social network to a point [20]

$$p_v = (x_v, y_v) \in \mathrm{R}^2,$$

the coordinates of *v* on the plot, is needed.

There are several desirable properties of an effective visualization: [20]

(1) Nodes should be separated by an optimal distance in order to fully utilize the two-dimensional space instead of being cluttered

(2) The length of a link should reflect the strength of association between the two end nodes, i.e., two connected nodes should appear closer if they are strongly associated, and distant if the association is weak.

(3) The crossing of edges should be minimized so the user can clearly see the relationships between nodes.

(4) The size of a node should be proportion to the importance of the corresponding terrorist.

(5) Degree of the actors

(6) Weighted Degree of the actors

(7) Eccentricity of the actors

(8) Closeness Centrality, Harmonic Closeness Centrality and Betweenness Centrality.

(9) Clustering Coefficient and Eigen Vector Centrality.

4. About the Terrorist Data Collected:

This project contains network analysis of the ISIS twitter dataset which was taken from:[2]

- **Data Description:**

They have scraped over 17,000 tweets from more than 100 posts of pro-ISIS fanboys from all over the world since the November 2015 Paris Attacks.

The dataset includes the following description of the data:

1. Name
2. Username
3. Description
4. Location
5. Number of followers at the time the tweet was downloaded
6. Number of statuses by the user when the tweet was downloaded
7. Date and timestamp of the tweet
8. The tweet itself

- **The top 5 details of the Data from the Dataset are:**

Out[64]:

	name	username	description	location	followers	numberstatuses	time	tweets
0	GunsandCoffee	GunsandCoffee70	ENGLISH TRANSLATIONS: http://t.co/QLdJ0ftews	NaN	640	49	1/6/2015 21:07	ENGLISH TRANSLATION: 'A MESSAGE TO THE TRUTHFU...
1	GunsandCoffee	GunsandCoffee70	ENGLISH TRANSLATIONS: http://t.co/QLdJ0ftews	NaN	640	49	1/6/2015 21:27	ENGLISH TRANSLATION: SHEIKH FATIH AL JAWLANI '...
2	GunsandCoffee	GunsandCoffee70	ENGLISH TRANSLATIONS: http://t.co/QLdJ0ftews	NaN	640	49	1/6/2015 21:29	ENGLISH TRANSLATION: FIRST AUDIO MEETING WITH ...
3	GunsandCoffee	GunsandCoffee70	ENGLISH TRANSLATIONS: http://t.co/QLdJ0ftews	NaN	640	49	1/6/2015 21:37	ENGLISH TRANSLATION: SHEIKH NASIR AL WUHAYSHI ...
4	GunsandCoffee	GunsandCoffee70	ENGLISH TRANSLATIONS: http://t.co/QLdJ0ftews	NaN	640	49	1/6/2015 21:45	ENGLISH TRANSLATION: AQAP: 'RESPONSE TO SHEIKH...

- There are total 17410 Rows in the dataset and 8 columns.

In [66]:
```
print(Tweet_data.shape)
```
```
(17410, 8)
```

- Which means that the total number of tweets in the dataset is: 17410
- But in this Dataset there are tweets which have been re-tweeted by the different users from a particular actor who is acting as the important actor in the network. The 1st picture shows the username(actors) and their respective tweets and the 2nd picture shows the username(actors) who have re-tweeted the original tweets of the different actors. Below is the top 5 sample from the dataset.

Out[79]:

	username	tweets
0	GunsandCoffee70	ENGLISH TRANSLATION: 'A MESSAGE TO THE TRUTHFU...
1	GunsandCoffee70	ENGLISH TRANSLATION: SHEIKH FATIH AL JAWLANI '...
2	GunsandCoffee70	ENGLISH TRANSLATION: FIRST AUDIO MEETING WITH ...
3	GunsandCoffee70	ENGLISH TRANSLATION: SHEIKH NASIR AL WUHAYSHI ...
4	GunsandCoffee70	ENGLISH TRANSLATION: AQAP: 'RESPONSE TO SHEIKH...

Unnamed: 0	username	tweets
0	GunsandCoffee70	RT @GIIMedia_CH004: Rules Of Imarah Part2 - Co...
1	abubakerdimshqi	RT @abdelhakzait: @nasseralfahad0كما فعل المجا...
2	abubakerdimshqi	RT @CNN: Photo of Israeli soldier holding Pale...
3	YazeedDhardaa25	RT @pieternanninga: New 37-min video #IS #Sina...
4	YazeedDhardaa25	RT @pieternanninga: New #IS video shows four m...

- Based on this a new dataset was created which describes the weight of each actor in the network along with the number of followers and number of statuses.

Out[56]:

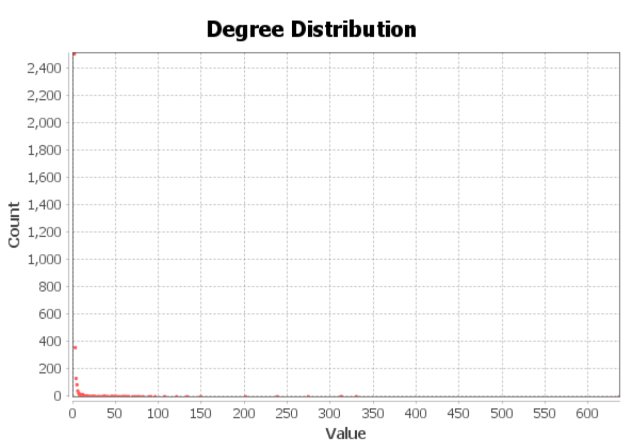
	Mentioned_statuses	Mentioned_followers	User	Mentions	Time	User_numberstatuses	User_followers	Weight
0	16688	29209	YazeedDhardaa25	RamiAlLolah	9/6/2015 20:58	127	904	1
1	16688	29209	YazeedDhardaa25	RamiAlLolah	9/8/2015 8:03	127	904	1
2	16688	29209	YazeedDhardaa25	RamiAlLolah	9/8/2015 8:04	127	904	1
3	16688	29209	YazeedDhardaa25	RamiAlLolah	9/9/2015 18:56	127	823	1
4	16688	29209	YazeedDhardaa25	RamiAlLolah	9/9/2015 18:56	127	823	1

5. Report of the Statistics of the Data:

Degree Report

Results:

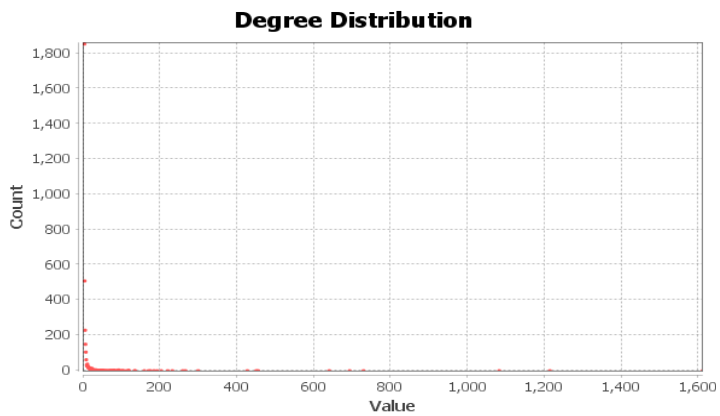
Average Degree: 3.204



Weighted Degree Report

Results:

Average Weighted Degree: 7.192



Graph Distance Report

Parameters:

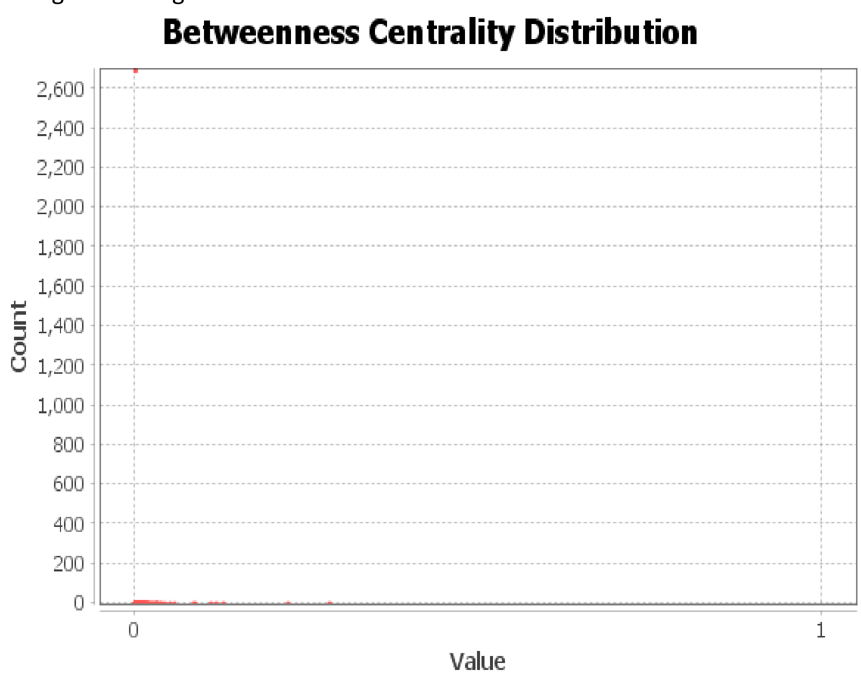
Network Interpretation: undirected

Results:

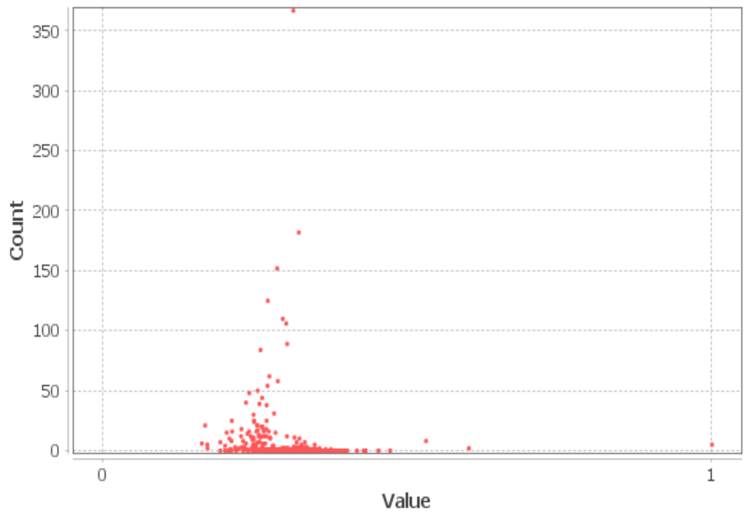
Diameter: 9

Radius: 1

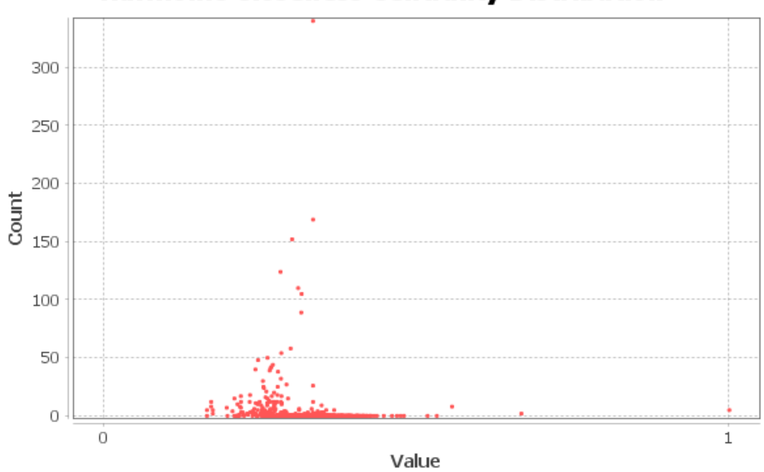
Average Path length: 3.597765924190278



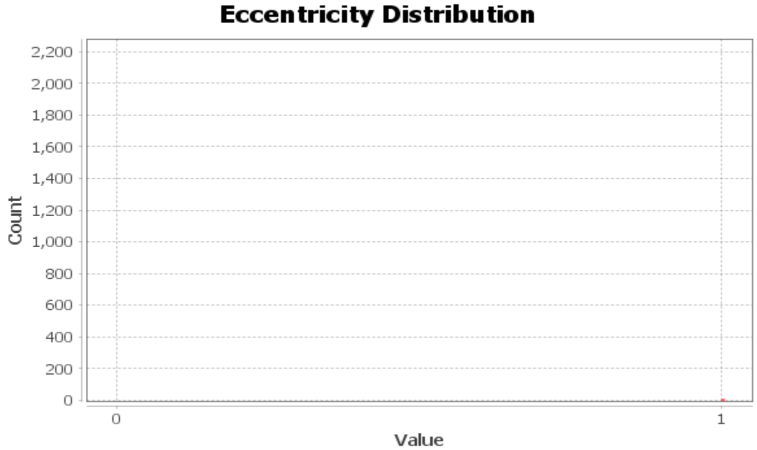
Closeness Centrality Distribution



Harmonic Closeness Centrality Distribution



Eccentricity Distribution



Graph Density Report

Parameters:

Network Interpretation: undirected

Results:

Density: 0.001

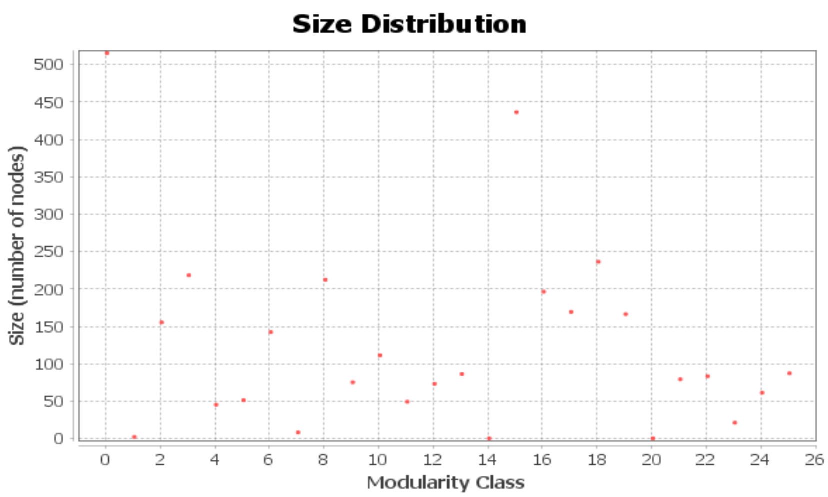
Modularity Report

Parameters:

Randomize: On
Use edge weights: On
Resolution: 1.0

Results:

Modularity: 0.649
Modularity with resolution: 0.649
Number of Communities: 26



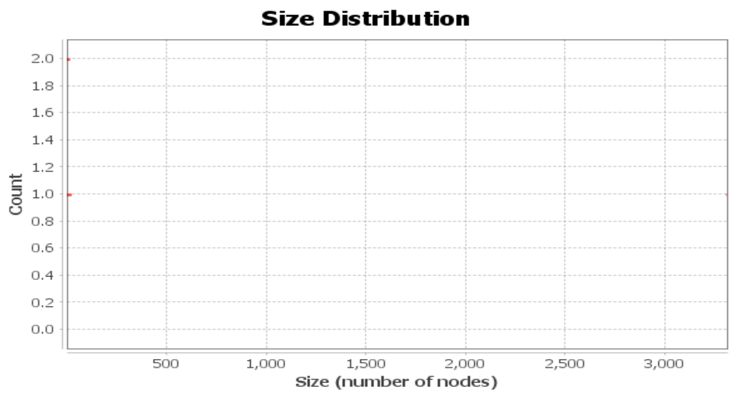
Connected Components Report

Parameters:

Network Interpretation: undirected

Results:

Number of Weakly Connected Components: 5



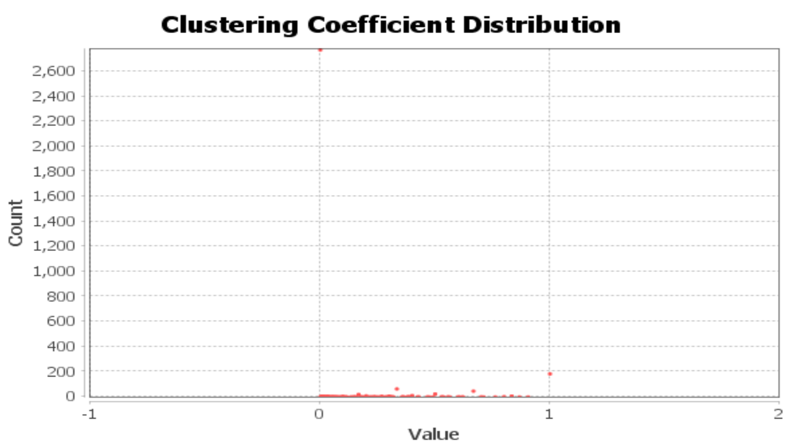
Clustering Coefficient Metric Report

Parameters:

Network Interpretation: undirected

Results:

Average Clustering Coefficient: 0.377
Total triangles: 2307
The Average Clustering Coefficient is the mean value of individual coefficients.

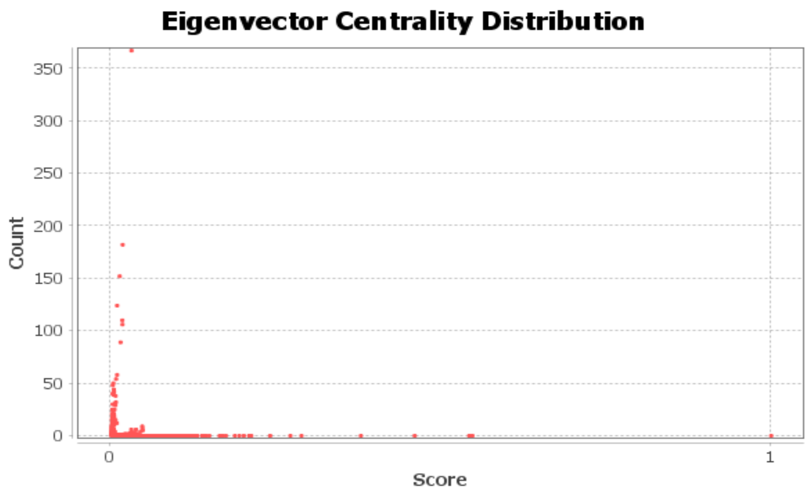


Eigenvector Centrality Report

Parameters:

Network Interpretation: undirected
Number of iterations: 100
Sum change: 0.030576453641535974

Results:

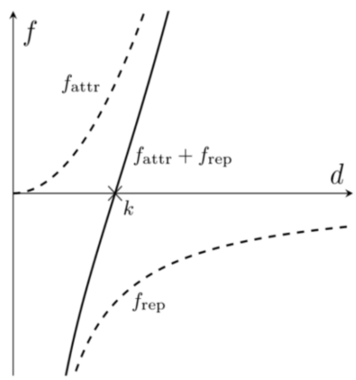


6. Computing the Network and Visualizing(Gephi Tool):

We have used the Fruchterman and Reingold Algorithm to initialize the co-ordinates of the given actors/nodes.
Some of the most flexible algorithms for calculating layouts of simple undirected graphs belong to a class known as force-directed algorithms. Also known as spring embedders, such algorithms calculate the layout of a graph using only information contained within the structure of the graph itself, rather than relying on domain-specific knowledge. Graphs drawn with these algorithms tend to be aesthetically pleasing, exhibit symmetries, and tend to produce crossing-free layouts for planar graphs. We employ a standard force-directed algorithm as a basis, that is, the algorithm of Fruchterman and Reingold. It assumes vertices to be point-shaped and defines two forces for influencing vertices: An attractive force f_{attr} that pulls connected vertices towards each other and a repulsive force f_{rep} that disperses the vertices by repelling them from each other. The absolute value of the forces can be computed as follows:[10][11][12]

- $f_{attr}(u, v) = k^2$ / distance(u, v)
- $f_{rep}(u, v) =$ distance$(u, v)^2$ / k

The directions of the forces are determined from the positions of the vertices given as two-dimensional vectors; for two vertices, the direction of repulsion and attraction is inverse. The complete force affecting a vertex v is computed by adding the repulsive forces for all other vertices and the attractive forces for all connected vertices together. As shown in the following figure, k describes the distance between two connected vertices whose attractive and repulsive forces are in equilibrium. [10][11][12]



The factor k is a constant and usually chosen according to the area of the drawing. If the distance between two vertices shrinks towards zero, the repulsive force grows infinitely. Similarly, for two connected vertices the attractive force grows with the distance between them. [10][11][12]

On this basis the force functions can be redefined. Let $d_{\text{actual}}(u, v)$ be the actual distance between two vertices u and v and let $d'_{X}(u, v)$ be the distance normalized by the minimum distance:

- $d_{\text{actual}}(u, v) = |\text{pos}(v) - \text{pos}(u)|$
- $d'_{X}(u, v) = d_{X}(u, v) - d_{\min}(u, v)$

This results in the following force functions, which comply with the criteria specified before:

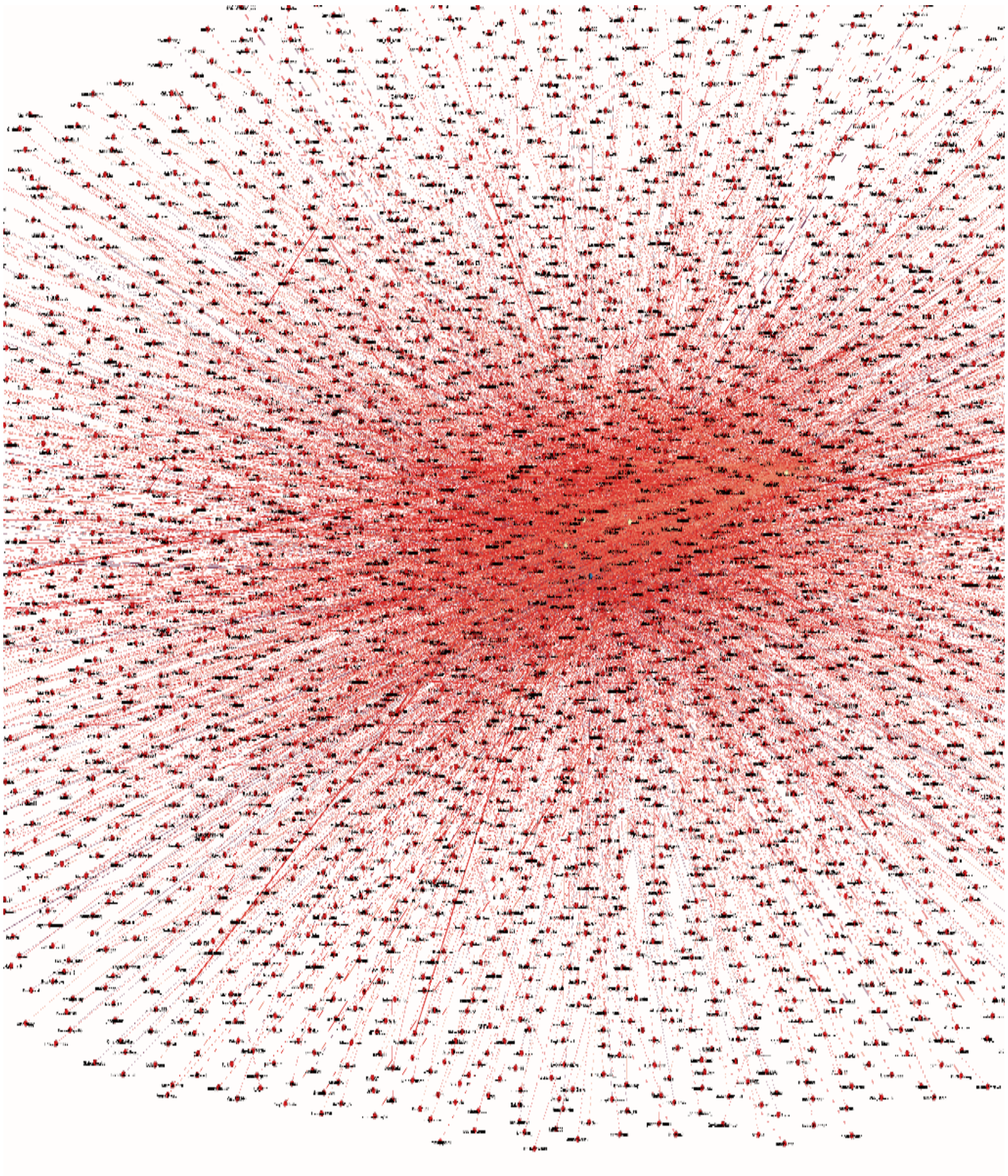
$$f_{\text{rep}}(r_u, r_v) = \begin{cases} \frac{d'_{\text{pref}}(r_u, r_v)^2}{d'_{\text{actual}}(r_u, r_v)} & \text{if } 0 \le d'_{\text{actual}}(r_u, r_v) \\ \infty & \text{otherwise} \end{cases}$$

$$f_{\text{attr}}(r_u, r_v) = \begin{cases} \frac{d'_{\text{actual}}(r_u, r_v)^2}{d'_{\text{pref}}(r_u, r_v)} & \text{if } 0 \le d'_{\text{actual}}(r_u, r_v) \\ 0 & \text{otherwise} \end{cases}$$

6.1 Computing the Nodes:

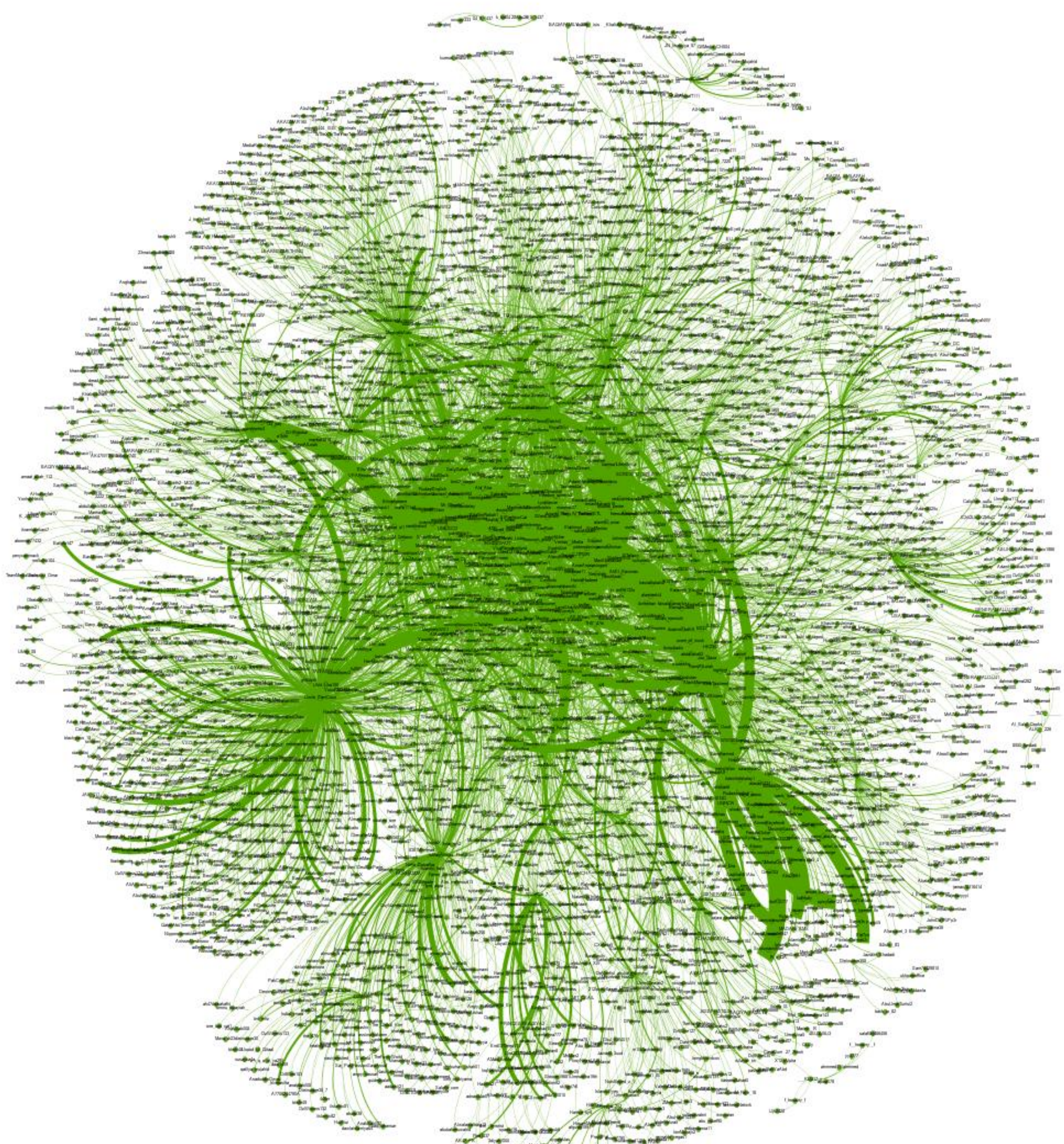
Each node v is displayed as a circle, whose size is controlled by its radius r_v. For the purpose of terrorist social network analysis, a node's prominence is largely determined by its centrality. In particular, we employed two centrality measures: degree and closeness. A node's degree $c_{degree}(v)$ is the number of links attached to it. An individual having a high degree may imply leadership while an individual with high closeness is more likely to serve as a mediator in the network. A node's closeness $c_{closeness}(v)$ is the inverse of the sum of its distances to all other nodes in the network,[20]

i.e., $c_{closeness}(v) = \dfrac{1}{\sum_{v \neq u \in V} \| p_u - p_v \|}$.

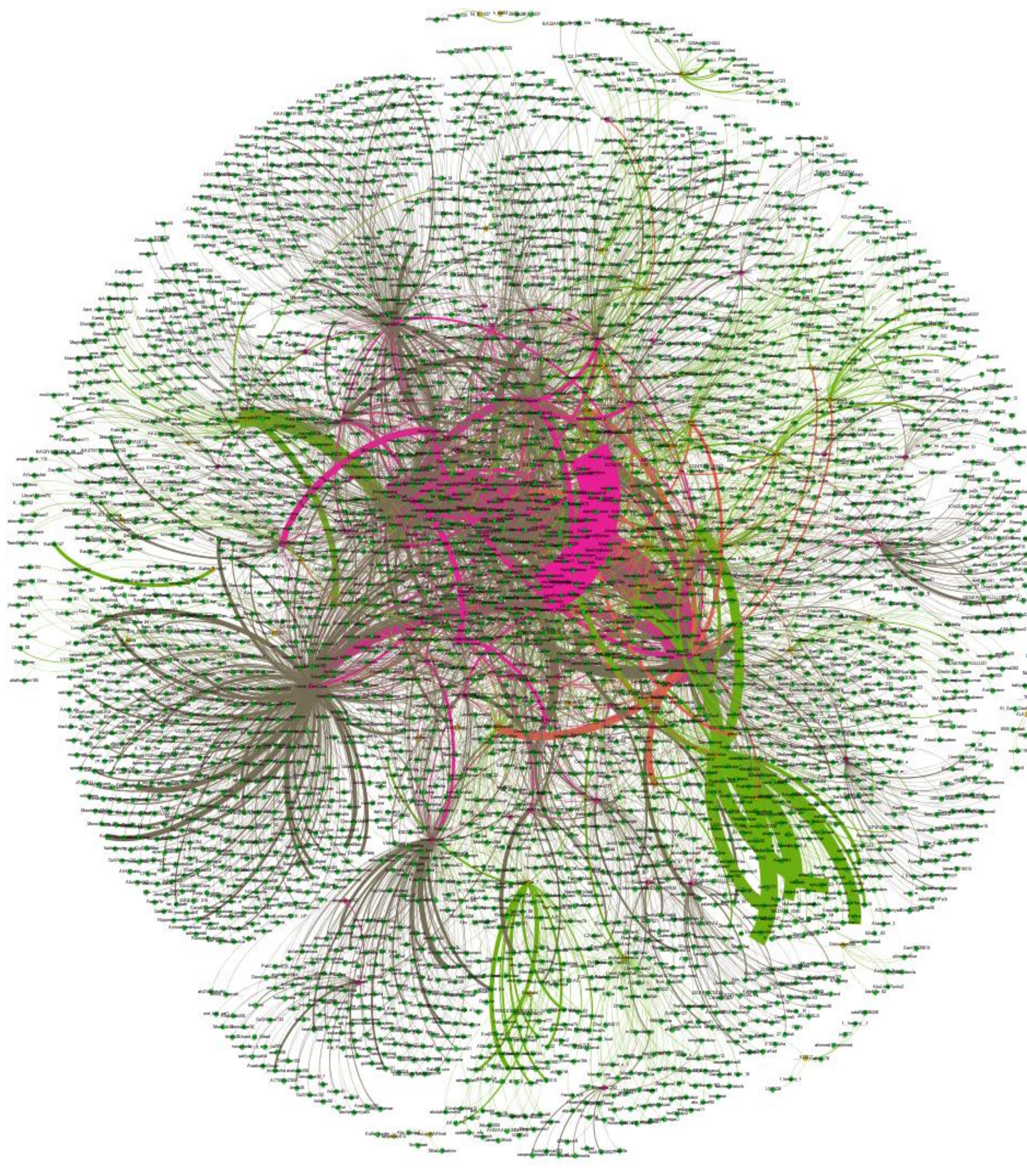


Before the Fruchterman and Reingold Algorithm was applied

After the Fruchterman and Reingold Algorithm was applied



- We have applied the Ranking of the weight to the edges to differentiate which edge has more weight and also the Nodes have been portioned by the different types of users : where “user” is the one who has done the actual tweet related to the ISIS and “mentions” are the users who have re-tweeted the tweets of those users and the third category where both “user and mention” are in the same tweet is shown with different nodes in the below diagram.

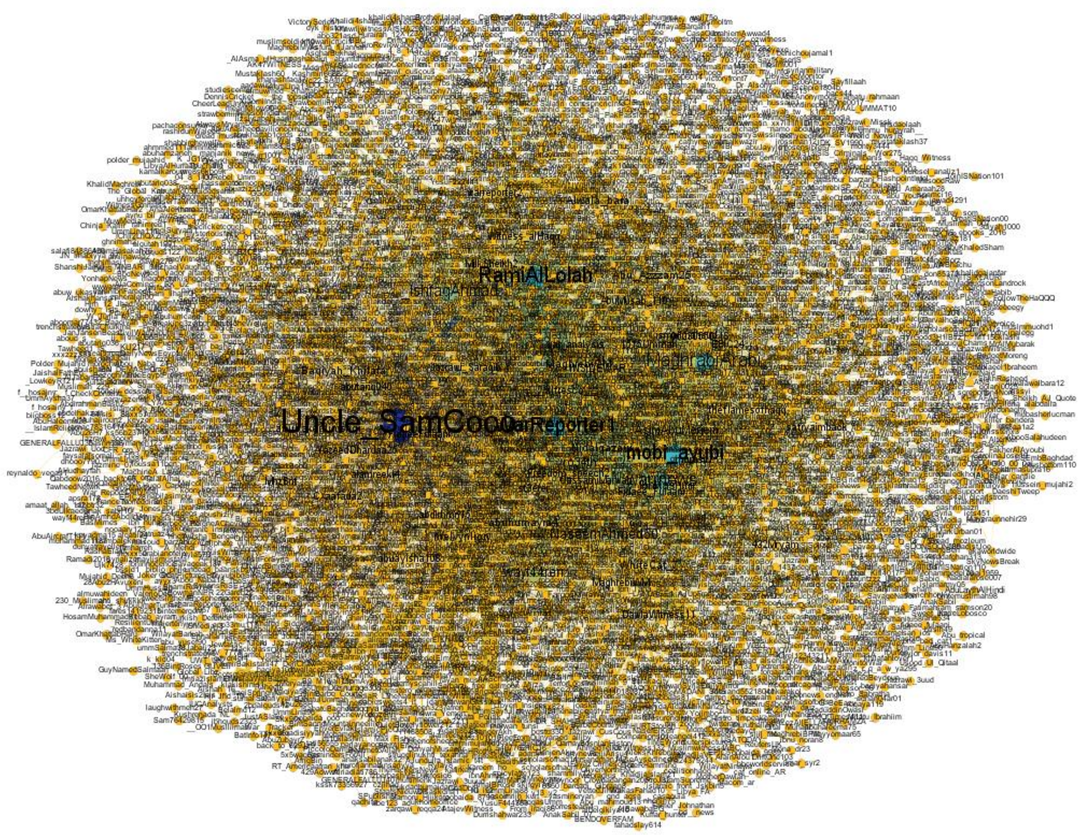


6.2 Measures Applied:

1. Degree:

The *degree* of a node, denoted by $d(n_i)$, is the number of lines that are incident with it. Equivalently, the degree of a node is the number of nodes adjacent to it. The degree of a node is a count that ranges from a minimum of 0, if no nodes are adjacent to a given node, to a maximum of g - 1, if a given node is adjacent to all other nodes in the graph. A node with degree equal to 0 is called an *isolate.* The degree of a node, $d(n_i)$, may be obtained by counting the number of lines incident with it.[16][21]

The Network generated is below:



From the above network based on the degree of the node we can see that the below tweeter users with the given usernames has the highest degree in the network and is connected to many nodes in the network. Based on this we can say that these users play an important role in the network to run this terrorist organization.

Data Table

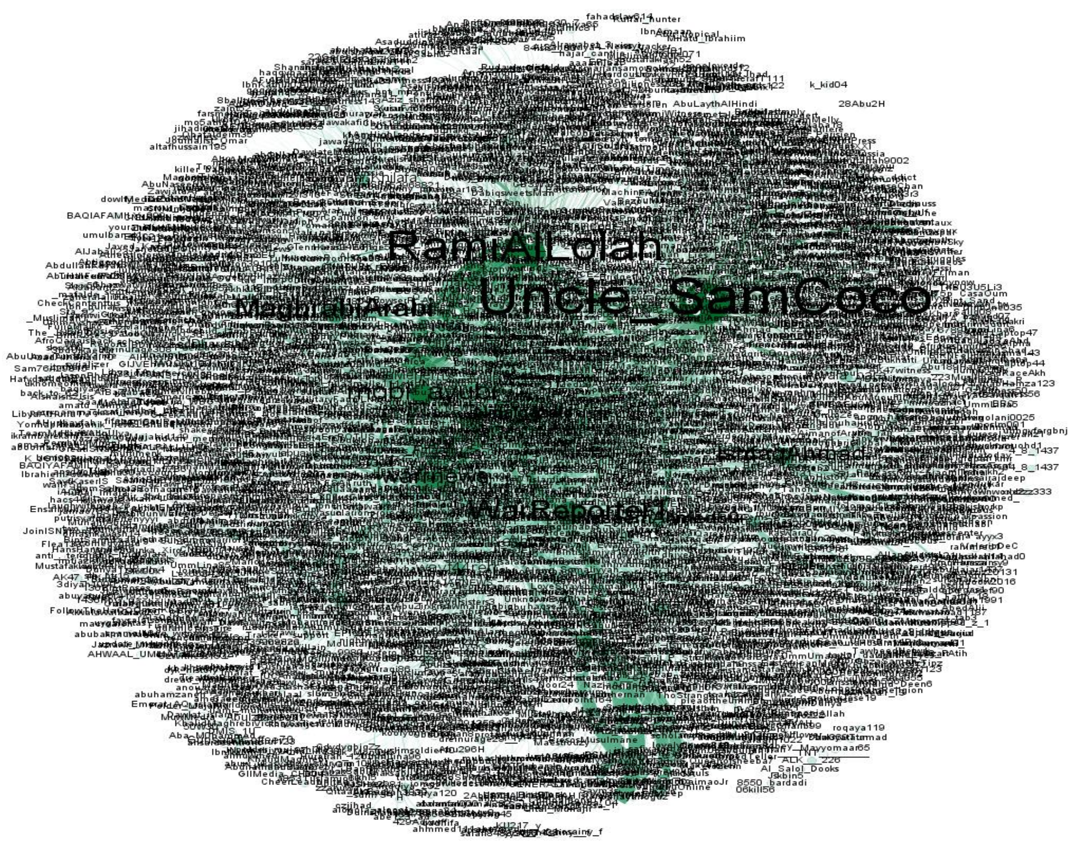
Nodes Edges Configuration Add node Add edge Search/Replace

Id	Label	Interval	frequency	type	Degree
Uncle_SamC…	Uncle_SamC…		1609	Mentions; User	635
RamiAlLolah	RamiAlLolah		1212	Mentions; User	329
mobi_ayubi	mobi_ayubi		1080	Mentions; User	312
WarReporter1	WarReporter1		727	Mentions; User	311
MaghrabiArabi	MaghrabiArabi		453	Mentions; User	273
warrnews	warrnews		691	Mentions; User	237
_IshfaqAhmad	_IshfaqAhmad		448	Mentions; User	200
NaseemAhm…	NaseemAhm…		257	Mentions; User	148
wayf44rerr	wayf44rerr		264	Mentions; User	132
IbnKashmir_	IbnKashmir_		218	; User	120

2. Weighted Degree:

In weighted networks the degree centrality is calculated as the sum of weights assigned to the node's direct connections and represents the node strength (Strength Centrality—SC). It is then based on tie weights and not on the number of ties. The disadvantage is that two nodes with the same strength, can be linked to a different number of nodes, and the initial information caught by DC is lost when SC is calculated. To overcome this disadvantage a tuning parameter has been defined to give relevance either to tie weights or number of ties alternatively. Hence, the weighted degree of a node is like the degree. It's based on the number of edge for a node, but pondered by the weight of each edge. It's doing the sum of the weight of the edges.[6][7][8][16]

The network generated by this is:



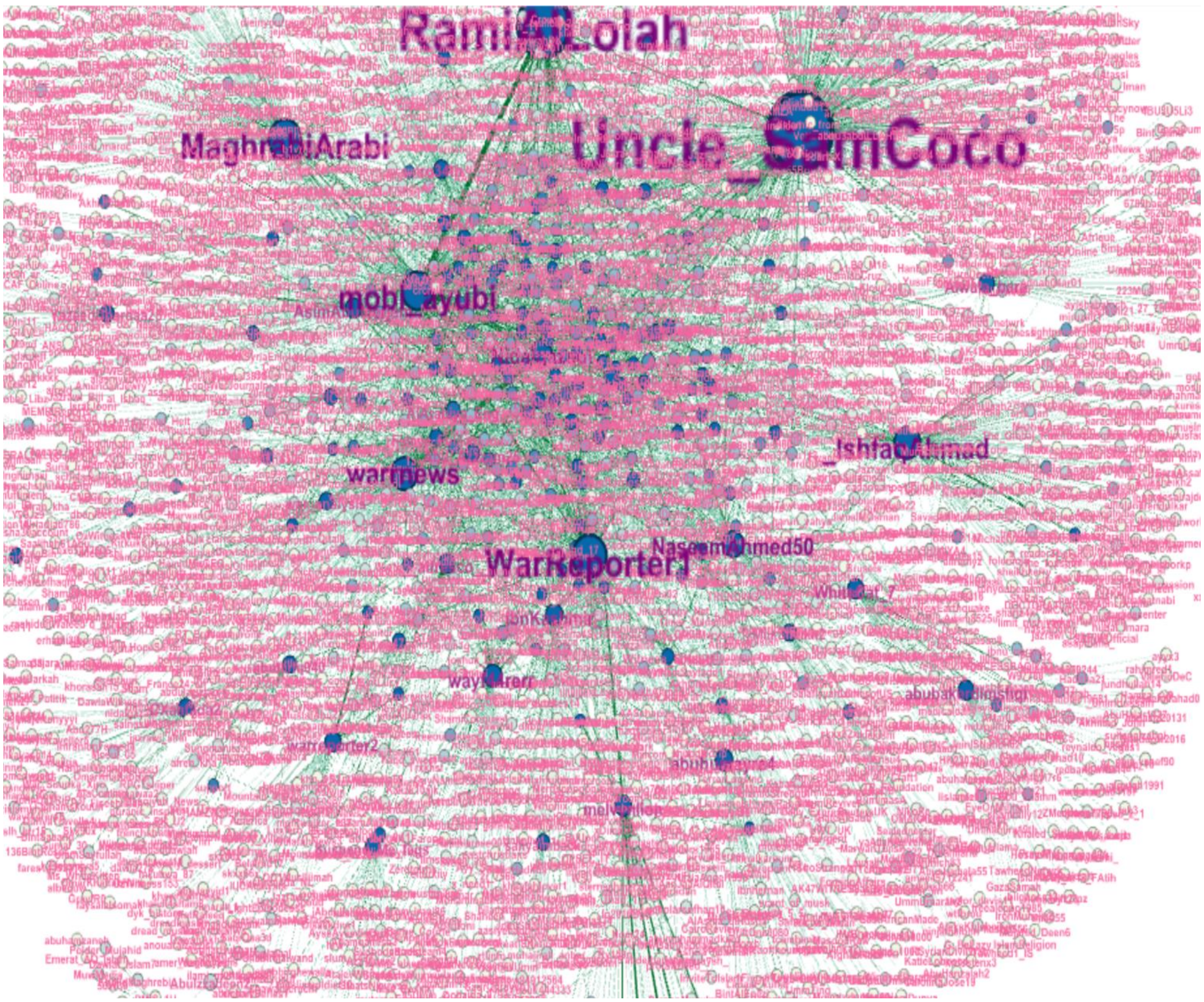
Based on the weighted degree of the network we can see that few of the actors who was having the highest degree remains the same, but few actors like after the actor “WarReporter1 “their positions have changed. Which means that these actors have more direct connections to the other nodes in the network and are responsible over how the information is flowing in the whole terrorist organization and these actors are mainly responsible for them.

Data Table

Nodes | Edges | Configuration | Add node | Add edge | Search/Replace | Import Spreadshe

Id	Label	Interval	frequency	type	Degree	Weighted Degr…
Uncle_SamC…	Uncle_SamC…		1609	Mentions; User	635	1609.0
RamiAlLolah	RamiAlLolah		1212	Mentions; User	329	1212.0
mobi_ayubi	mobi_ayubi		1080	Mentions; User	312	1080.0
WarReporter1	WarReporter1		727	Mentions; User	311	727.0
warrnews	warrnews		691	Mentions; User	237	691.0
melvynlion	melvynlion		638	; User	77	638.0
MaghrabiArabi	MaghrabiArabi		453	Mentions; User	273	453.0
_IshfaqAhmad	_IshfaqAhmad		448	Mentions; User	200	448.0
Nidalgazaui	Nidalgazaui		425	Mentions; User	81	425.0
AsimAbuMer…	AsimAbuMer…		297	Mentions; User	106	297.0

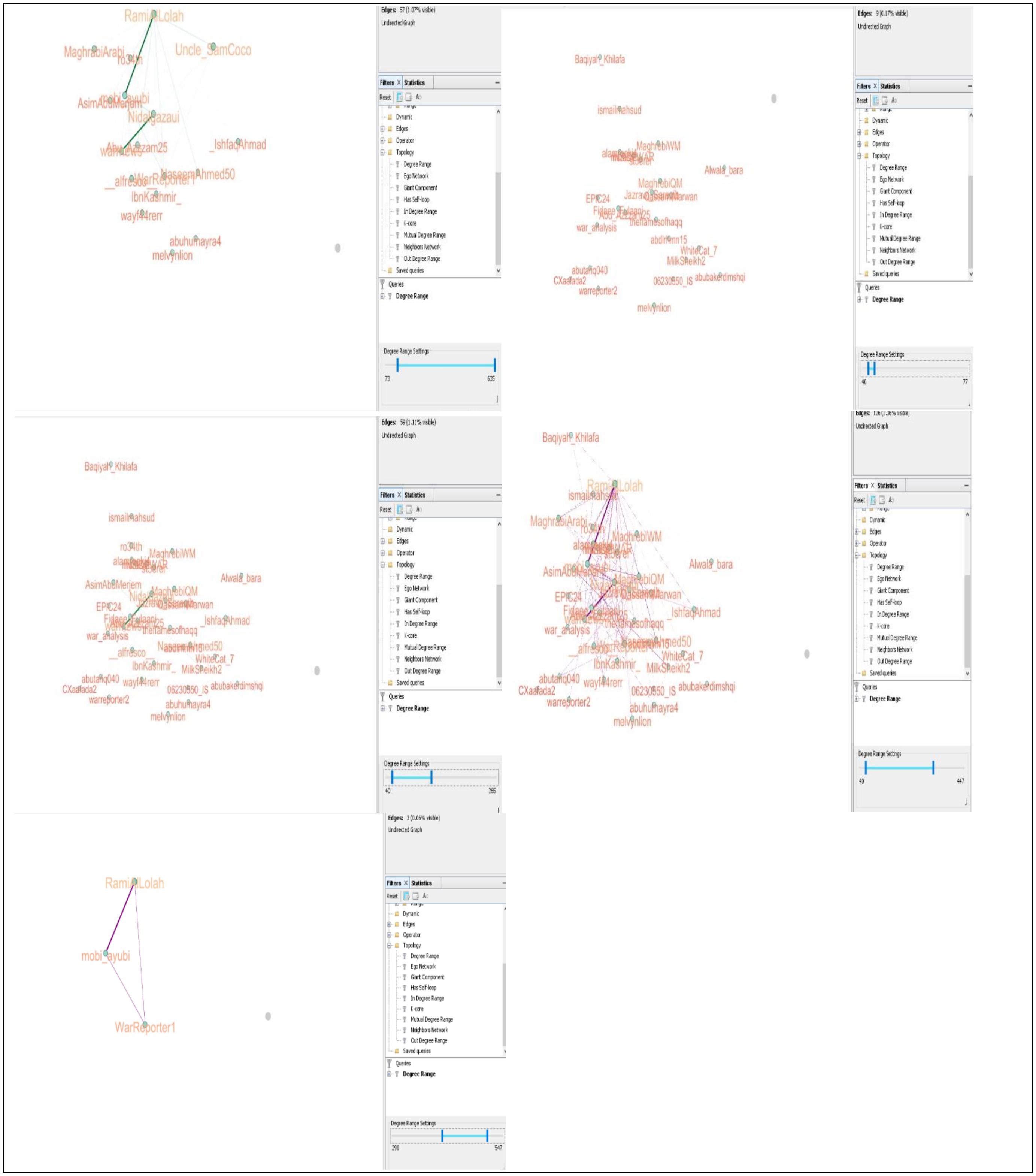
When we apply the Fruchterman and Reingold Algorithm further on this network we can see that the network breaks up into more visible clusters as below:



The below network structures show the important actors in the terrorist organization who are the bridge between all the other nodes and these actors take charge of the tweets in the twitter to spread terrorism and influence people through their tweets.

3. Degree Range:

When we sort out the network to view more distinguished network in-order to identify the actors which are grouped together or are having direct connections or ties with each other depending on the degree range variation it helps us to visualization more about the important actors in the terrorist organization and how these actors are influencing more other actors in the network and their connectivity as the range changes, which says about how many followers these actors are having in the network and the most influential and least influential actors among them.[6][7][8]



RamiALolah
MaghrebiArabi
ro34th
Uncle_SamCoco
AsimAbuMerjem
mobi_ayubi
NidalGazaui
_IshfaqAhmad
Abu_Azzam25
Naseem_Ahmed50
__alfresco__
WarReporter1
IbnKashmir_
wayf44rerr
abuhumayra4
melvynlion
Baqiyah_Khilafa
ismailmahsud
MaghrebiWM
Alwala_bara
MaghrebiQM
EPIC24
war_analysis
abdirnmn15
WhiteCat_7
MilkSheikh2
abutariq040
CXaafada2
warreporter2
06230550_IS
abubakerdimshqi
theflameofhaqq
Edges:
Undirected Graph
Filters
Statistics
Reset
Dynamic
Edges
Operator
Topology
Degree Range
Ego Network
Giant Component
Has Self-loop
In Degree Range
K-core
Mutual Degree Range
Neighbors Network
Out Degree Range
Saved queries
Queries
Degree Range
Degree Range Settings

4. K-core:

In graph theory, a k-degenerate graph is an undirected graph in which every subgraph has a vertex of degree at most k: that is, some vertex in the subgraph touches k or fewer of the subgraph's edges. The degeneracy of a graph is the smallest value of k for which it is k-degenerate. The degeneracy of a graph is a measure of how sparse it is and is within a constant factor of other sparsity measures such as the arboricity of a graph.[5]

- Degeneracy is also known as the k-core number, width, and linkage, and is essentially the same as the coloring number or Szekeres-Wilf number (named after Szekeres and Wilf (1968)). [5]
- k-degenerate graphs have also been called k-inductive graphs. [5]
- The degeneracy of a graph may be computed in linear time by an algorithm that repeatedly removes minimum-degree vertices. [5]
- The connected components that are left after all vertices of degree less than k have been removed are called the k-cores of the graph and the degeneracy of a graph is the largest value k such that it has a k-core. [5]
- A k-core of a graph G is a maximal connected subgraph of G in which all vertices have degree at least k. [5]
- Equivalently, it is one of the connected components of the subgraph of G formed by repeatedly deleting all vertices of degree less than k. If a non-empty k-core exists, then, clearly, G has degeneracy at least k, and the degeneracy of G is the largest k for which G has a k-core. [5]
- A vertex u has coreness if it belongs to c-core but not to any (c+1)-core. [5]
- The concept of a k-core was introduced to study the clustering structure of social networks and to describe the evolution of random graphs. [5]

In more detail, the algorithm proceeds as follows: [5]

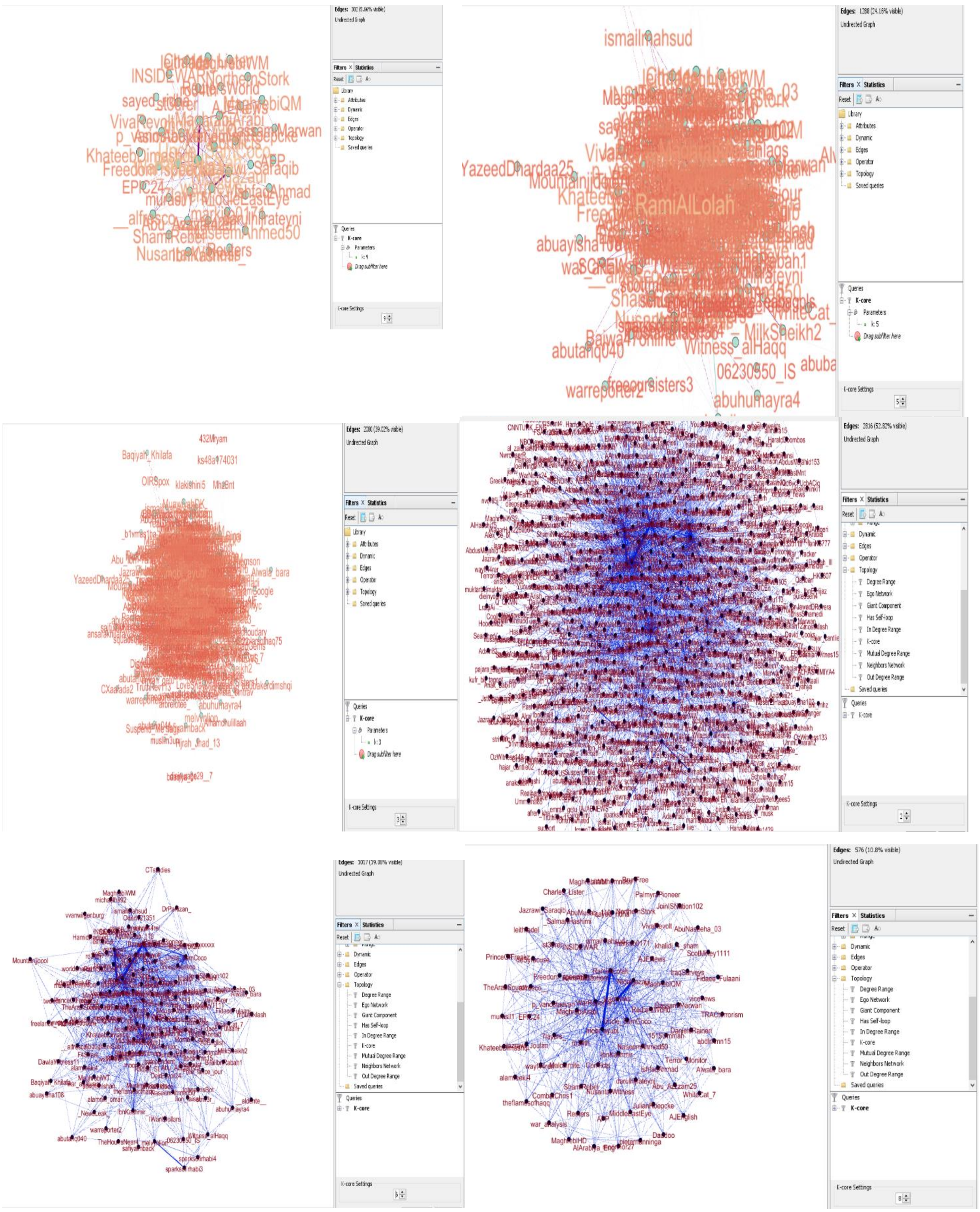
- Initialize an output list *L*. [5]
- Compute a number d_v for each vertex *v* in *G*, the number of neighbors of *v* that are not already in *L*. Initially, these numbers are just the degrees of the vertices. [5]
- Initialize an array *D* such that *D*[*i*] contains a list of the vertices *v* that are not already in *L* for which $d_v = i$. [5]
- Initialize *k* to 0. [5]
- Repeat *n* times: [5]
  - Scan the array cells *D*[0], *D*[1], ... until finding an *i* for which *D*[*i*] is nonempty. [5]
  - Set *k* to max(*k*,*i*) [5]
  - Select a vertex *v* from *D*[*i*]. Add *v* to the beginning of *L* and remove it from *D*[*i*]. [5]
  - For each neighbor *w* of *v* not already in *L*, subtract one from d_w and move *w* to the cell of D corresponding to the new value of d_w. [5]

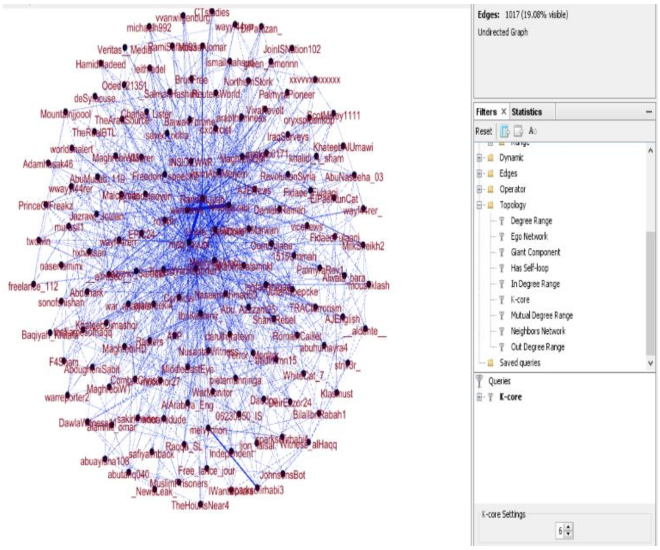
At the end of the algorithm, *k* contains the degeneracy of *G* and *L* contains a list of vertices in an optimal ordering for the coloring number. The *i*-cores of *G* are the prefixes of *L* consisting of the vertices added to *L* after *k* first takes a value greater than or equal to *i*. [5]

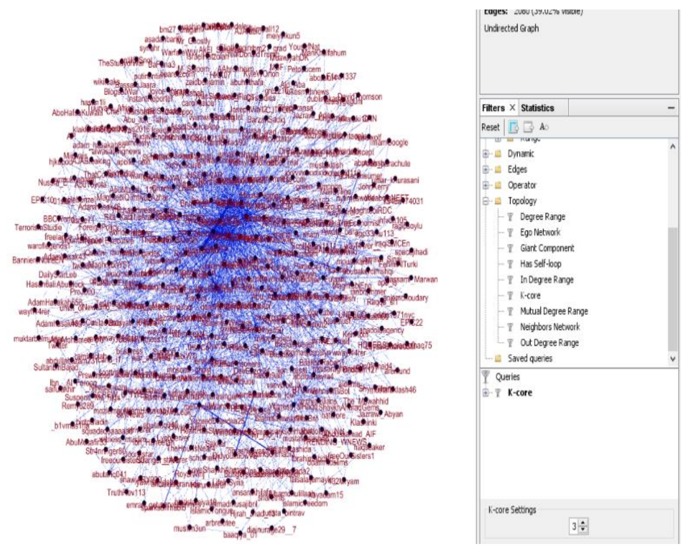
Initializing the variables *L*, d_v, *D*, and *k* can easily be done in linear time. Finding each successively removed vertex *v* and adjusting the cells of *D* containing the neighbors of *v* take time proportional to the value of d_v at that step; but the sum of these values is the number of edges of the graph (each edge contributes to the term in the sum for the later of its two endpoints) so the total time is linear. [5]

Hence, when we apply this K-core in the terrorist organization graph we can see all the actors that are related to each other after all the actors with the degree less than the k has been removed and the degeneracy of the graph is generated. It also helps us in understanding the actors which are always communicating with each other and they form a thick cluster together. Which means that any information received to one will be transmitted to all other actors in that k-core group and the information keeps flowing for the terrorists. [5]

Below is the network visualization of the different k-core values:







5. Clustering Coefficient:

Clustering coefficient is a measure of the degree to which nodes in a graph tend to cluster together. Evidence suggests that in most real-world networks, and in particular social networks, nodes tend to create tightly knit groups characterized by a relatively high density of ties; this likelihood tends to be greater than the average probability of a tie randomly established between two nodes. Two versions of this measure exist: the global and the local. The global version was designed to give an overall indication of the clustering in the network, whereas the local gives an indication of the embeddedness of single nodes. The global clustering coefficient is based on triplets of nodes. A triplet is three nodes that are connected by either two (open triplet) or three (closed triplet) undirected ties. A triangle graph therefore includes three closed triplets, one centered on each of the nodes (n.b. this means the three triplets in a triangle come from overlapping selections of nodes). The global clustering coefficient is the number of closed triplets (or 3 x triangles) over the total number of triplets (both open and closed). The first attempt to measure it was made by Luce and Perry (1949). This measure gives an indication of the clustering in the whole network (global) and can be applied to both undirected and directed networks (often called transitivity, see Wasserman and Faust, 1994, page 243[4]).

The global clustering coefficient is defined as:

$$C = \frac{\text{number of closed triplets}}{\text{number of all triplets (open and closed)}}.$$

The number of closed triplets has also been referred to as 3 × triangles in the literature, so:

$$C = \frac{3 \times \text{number of triangles}}{\text{number of all triplets}}.$$

The **local clustering coefficient** of a vertex (node) in a graph quantifies how close its neighbours are to being a clique (complete graph). Duncan J. Watts and Steven Strogatz introduced the measure in 1998 to determine whether a graph is a small-world network.

A graph $G = (V, E)$ formally consists of a set of vertices V and a set of edges E between them. An edge e_{ij} connects vertex v_i with vertex v_j.

The neighbourhood N_i for a vertex v_i is defined as its immediately connected neighbours as follows:

$$N_i = \{v_j : e_{ij} \in E \vee e_{ji} \in E\}.$$

We define k_i as the number of vertices, $|N_i|$, in the neighbourhood, N_i, of a vertex.

The local clustering coefficient C_i for a vertex v_i is then given by the proportion of links between the vertices within its neighbourhood divided by the number of links that could possibly exist between them. For a directed graph, e_{ij} is distinct from e_{ji}, and therefore for each neighbourhood N_i there are $k_i(k_i - 1)$ links that could exist among the vertices within the neighbourhood (k_i is the number of neighbours of a vertex). Thus, the **local clustering coefficient for directed graphs** is given as [2]

$$C_i = \frac{|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)}.$$

An undirected graph has the property that e_{ij} and e_{ji} are considered identical. Therefore, if a vertex v_i has k_i neighbours, $\frac{k_i(k_i-1)}{2}$ edges could exist among the vertices within the neighbourhood. Thus, the **local clustering coefficient for undirected graphs** can be defined as

$$C_i = \frac{2|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)}.$$

Let $\lambda_G(v)$ be the number of triangles on $v \in V(G)$ for undirected graph G. That is, $\lambda_G(v)$ is the number of subgraphs of G with 3 edges and 3 vertices, one of which is v. Let $\tau_G(v)$ be the number of triples on $v \in G$. That is, $\tau_G(v)$ is the number of subgraphs (not necessarily induced) with 2 edges and 3 vertices, one of which is v and such that v is incident to both edges. Then we can also define the clustering coefficient as

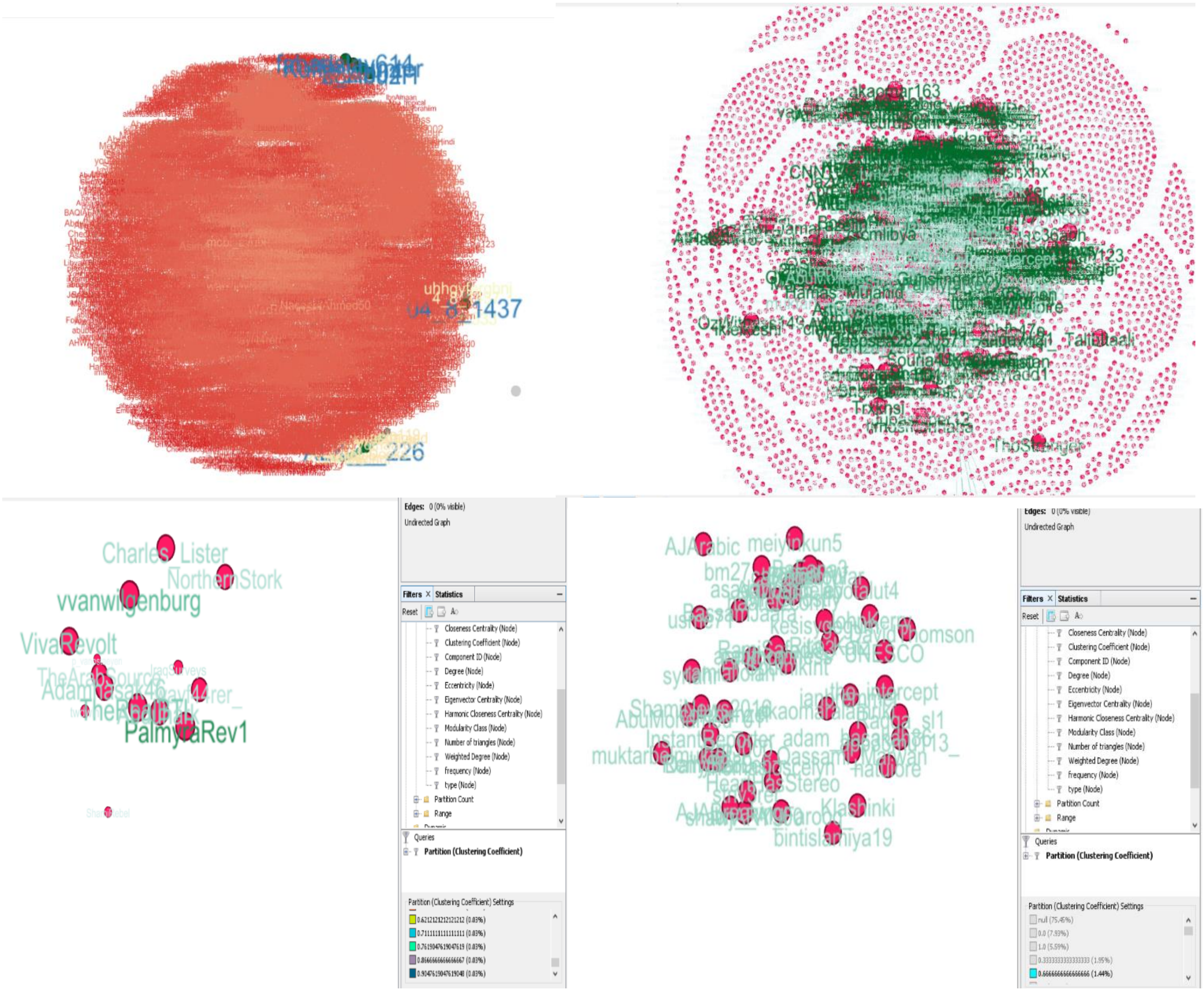
$$C_i = \frac{\lambda_G(v)}{\tau_G(v)}.$$

It is simple to show that the two preceding definitions are the same, since

$$\tau_G(v) = C(k_i, 2) = \frac{1}{2} k_i(k_i - 1).$$

These measures are 1 if every neighbour connected to v_i is also connected to every other vertex within the neighbourhood, and 0 if no vertex that is connected to v_i connects to any other vertex that is connected to v_i.

Below is the network visualization of the different Clustering coefficient values:



Based on this we have identified the below actors who are having high clustering coefficient value and who is responsible for maintaining the closed groups in this terrorist organization and also based on these clusters the various groups performs the various activities that have been assigned to them to maintain this terrorist organization.

Data Table

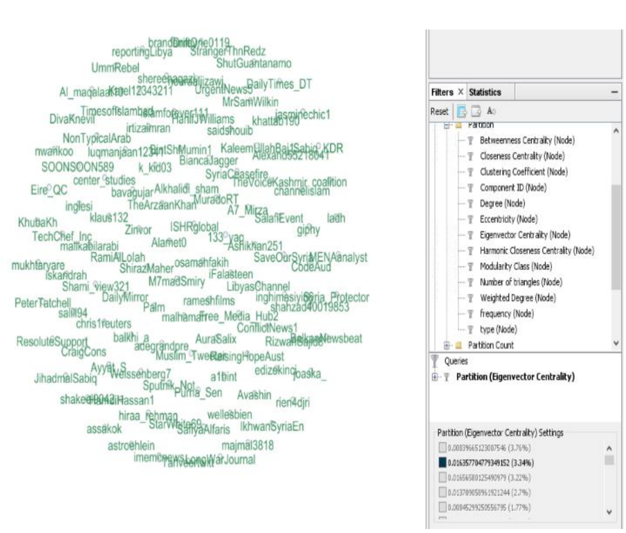
Nodes Edges Configuration Add node Add edge Search/Replace Import Spreadsheet Export table More actions Filter: Id

Id	Label	Interval	frequency	type	Degree	Weighted Degr…	Eccentricity	Closeness Centrality	Harmonic Closeness Centrality	Betweenness Centrality	Modularity Class	Component ID	Clustering Coefficient	Number of triangles
Uncle_SamC…	Uncle_SamC…		1609	Mentions; User	635	1609.0	5.0	0.451186	0.531157	0.28261	12	0	0.003184	641
WarReporter1	WarReporter1		727	Mentions; User	311	727.0	5.0	0.426803	0.478523	0.127907	15	0	0.010497	506
RamiAlLolah	RamiAlLolah		1212	Mentions; User	329	1212.0	5.0	0.470296	0.51639	0.222185	2	0	0.007988	431
mobi_ayubi	mobi_ayubi		1080	Mentions; User	312	1080.0	5.0	0.415704	0.467362	0.109905	24	0	0.008863	430
warrnews	warrnews		691	Mentions; User	237	691.0	5.0	0.429964	0.473486	0.087075	24	0	0.015054	421
Nidalgazaui	Nidalgazaui		425	Mentions; User	81	425.0	5.0	0.429461	0.459363	0.050321	24	0	0.079321	257
MaghrabiArabi	MaghrabiArabi		453	Mentions; User	273	453.0	5.0	0.397382	0.446127	0.117879	0	0	0.004956	184
NaseemAhm…	NaseemAhm…		257	Mentions; User	148	257.0	5.0	0.399541	0.435207	0.05702	25	0	0.015076	164
MaghrebiQM	MaghrebiQM		168	Mentions; User	72	168.0	5.0	0.392062	0.423824	0.016718	6	0	0.057121	146
__alfresco__	__alfresco__		230	Mentions; User	88	230.0	5.0	0.387289	0.419316	0.028267	15	0	0.035789	137
wayf44rerr	wayf44rerr		264	Mentions; User	132	264.0	6.0	0.367096	0.40621	0.042672	15	0	0.01307	113
IbnKashmir_	IbnKashmir_		218	; User	120	218.0	5.0	0.372719	0.40277	0.036962	3	0	0.015126	108

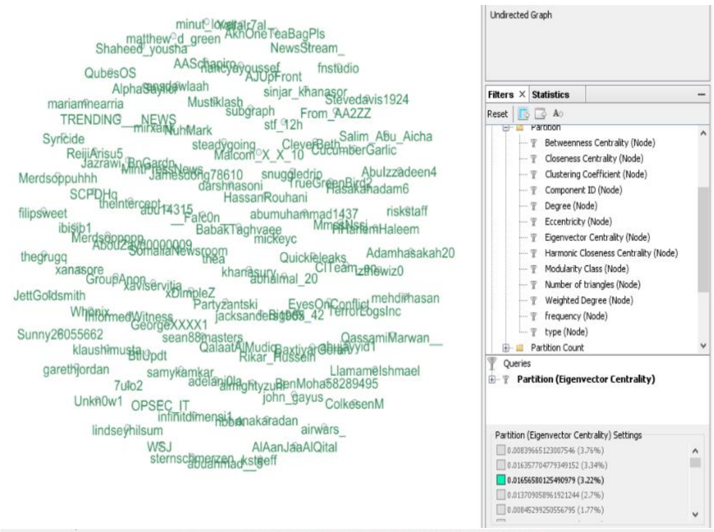
6. Eigenvector Centrality:

Eigenvector centrality (also called eigencentrality or prestige score) is a measure of the influence of a node in a network. Relative scores are assigned to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. A high eigenvector score means that a node is connected to many nodes who themselves have high scores.[13]

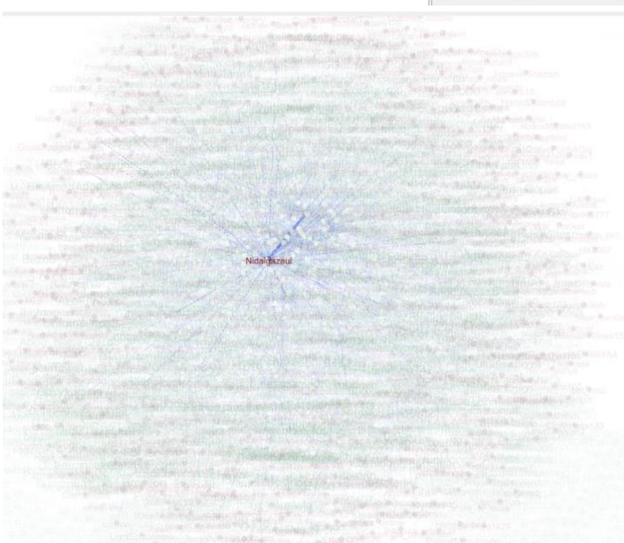
Below is the network with Eigen Vector Centrality:



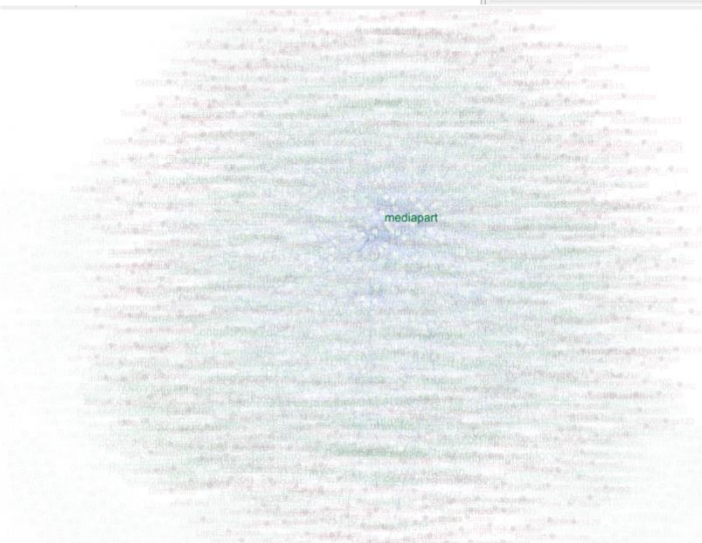
Filters
Statistics
Reset
Partition
Betweenness Centrality (Node)
Closeness Centrality (Node)
Clustering Coefficient (Node)
Component ID (Node)
Degree (Node)
Eccentricity (Node)
Eigenvector Centrality (Node)
Harmonic Closeness Centrality (Node)
Modularity Class (Node)
Number of triangles (Node)
Weighted Degree (Node)
frequency (Node)
type (Node)
Partition Count
Queries
Partition (Eigenvector Centrality)
Partition (Eigenvector Centrality) Settings



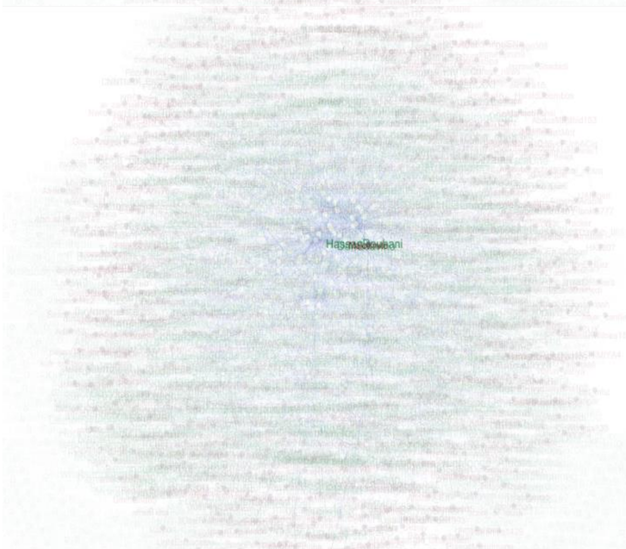
Undirected Graph
Filters
Statistics
Reset
Partition
Betweenness Centrality (Node)
Closeness Centrality (Node)
Clustering Coefficient (Node)
Component ID (Node)
Degree (Node)
Eccentricity (Node)
Eigenvector Centrality (Node)
Harmonic Closeness Centrality (Node)
Modularity Class (Node)
Number of triangles (Node)
Weighted Degree (Node)
frequency (Node)
type (Node)
Partition Count
Queries
Partition (Eigenvector Centrality)
Partition (Eigenvector Centrality) Settings

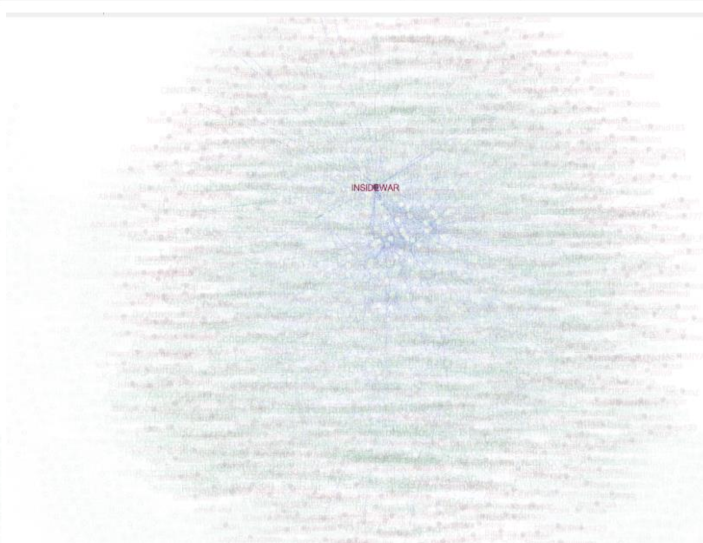


Nidalgazaui



mediapart





INSIDEWAR

Below is the description of the most influential actors of this whole terrorist organization who servers as the backbone in the digital world to spread hatred, ISIS related messages and who also controls the digital group of this terrorist organization and are the most dangerous one.

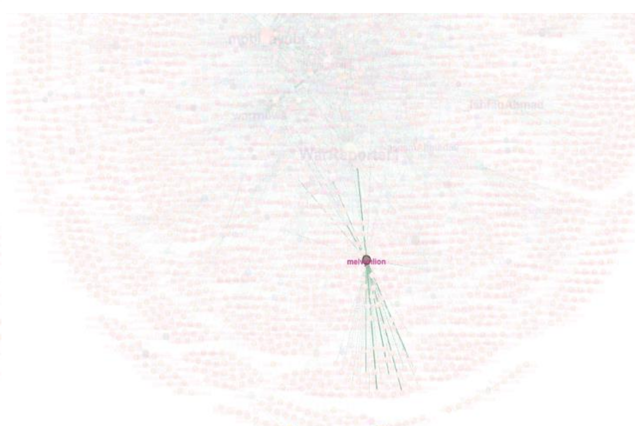
Id	Label	Interval	frequency	type	Degree	Weighted Degr…	Eccentricity	Closeness Centrality	Harmonic Closeness Centrality	Betweenness Centrality	Modularity Class	Component ID	Clustering Coefficient	Number of triangles	Eigenvector Centrality
Uncle_SamC…	Uncle_SamC…		1609	Mentions; User	635	1609.0	5.0	0.451186	0.531157	0.28261	12	0	0.003184	641	1.0
WarReporter1	WarReporter1		727	Mentions; User	311	727.0	5.0	0.426803	0.478523	0.127907	15	0	0.010497	506	0.547016
RamiAlLolah	RamiAlLolah		1212	Mentions; User	329	1212.0	5.0	0.470296	0.51639	0.222185	2	0	0.007988	431	0.542752
mobi_ayubi	mobi_ayubi		1080	Mentions; User	312	1080.0	5.0	0.415704	0.467362	0.109905	24	0	0.008863	430	0.542294
warrnews	warrnews		691	Mentions; User	237	691.0	5.0	0.429964	0.473486	0.087075	24	0	0.015054	421	0.459484
MaghrabiArabi	MaghrabiArabi		453	Mentions; User	273	453.0	5.0	0.397382	0.446127	0.117879	0	0	0.004956	184	0.378075
Nidalgazaui	Nidalgazaui		425	Mentions; User	81	425.0	5.0	0.429461	0.459363	0.050321	24	0	0.079321	257	0.287859
NaseemAhm…	NaseemAhm…		257	Mentions; User	148	257.0	5.0	0.399541	0.435207	0.05702	25	0	0.015076	164	0.271118
_IshfaqAhmad	_IshfaqAhmad		448	Mentions; User	200	448.0	5.0	0.368158	0.405465	0.084707	14	0	0.003367	67	0.240442
MaghrebiQM	MaghrebiQM		168	Mentions; User	72	168.0	5.0	0.392062	0.423824	0.016718	6	0	0.057121	146	0.211763
wayf44rerr	wayf44rerr		264	Mentions; User	132	264.0	6.0	0.367096	0.40621	0.042672	15	0	0.01307	113	0.209822
IbnKashmir_	IbnKashmir_		218	; User	120	218.0	5.0	0.372719	0.40277	0.036962	3	0	0.015126	108	0.208721

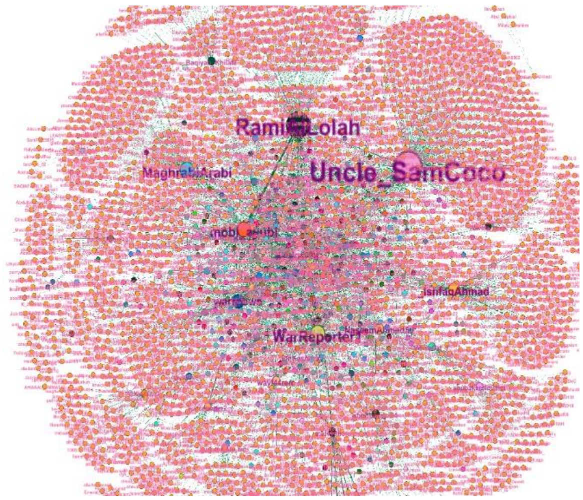
7. Closeness Centrality, Harmonic Closeness Centrality and Betweenness Centrality:

Based on the closeness Centrality, Harmonic closeness Centrality and Betweenness Centrality we can see the actors which form the central node of the complete Network. These actors below also serve as a bridge between the other actors in this terrorist Network. Which means these are the most important actors of the terrorist network organization which helps in connecting and information flow between them in the complete network. These actors help in spreading information which is being generated by the most wanted terrorist users from the network.[21]

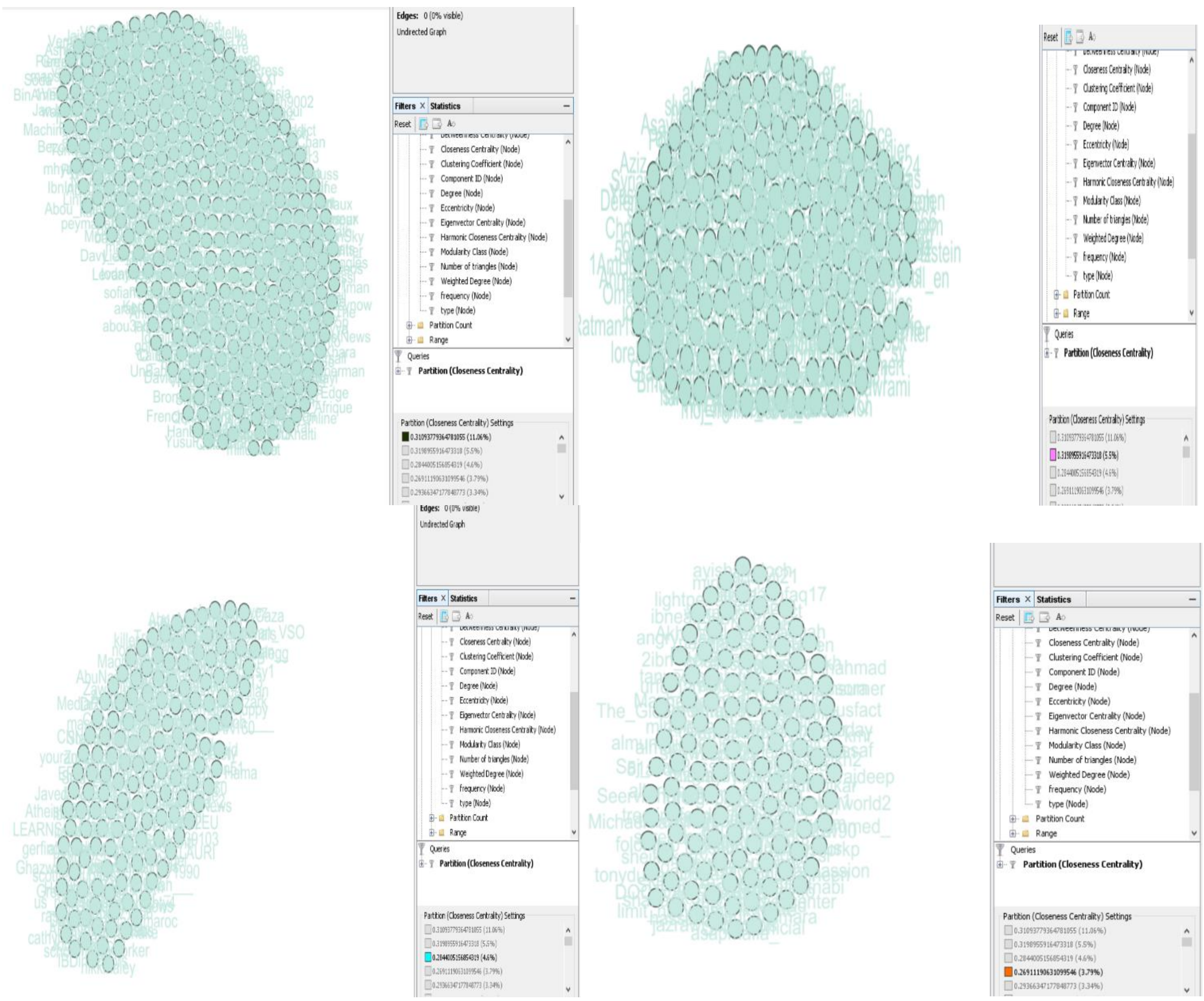
Id	Label	Interval	frequency	type	Degree	Weighted Degr…	Eccentricity	Closeness Centrality	Harmonic Closeness Centrality	Betweenness Centrality
Uncle_SamC…	Uncle_SamC…		1609	Mentions; User	635	1609.0	5.0	0.451186	0.531157	0.28261
RamiAlLolah	RamiAlLolah		1212	Mentions; User	329	1212.0	5.0	0.470296	0.51639	0.222185
WarReporter1	WarReporter1		727	Mentions; User	311	727.0	5.0	0.426803	0.478523	0.127907
MaghrabiArabi	MaghrabiArabi		453	Mentions; User	273	453.0	5.0	0.397382	0.446127	0.117879
mobi_ayubi	mobi_ayubi		1080	Mentions; User	312	1080.0	5.0	0.415704	0.467362	0.109905
warrnews	warrnews		691	Mentions; User	237	691.0	5.0	0.429964	0.473486	0.087075
_IshfaqAhmad	_IshfaqAhmad		448	Mentions; User	200	448.0	5.0	0.368158	0.405465	0.084707
NaseemAhm…	NaseemAhm…		257	Mentions; User	148	257.0	5.0	0.399541	0.435207	0.05702
Nidalgazaui	Nidalgazaui		425	Mentions; User	81	425.0	5.0	0.429461	0.459363	0.050321
wayf44rerr	wayf44rerr		264	Mentions; User	132	264.0	6.0	0.367096	0.40621	0.042672
abuhumayra4	abuhumayra4		157	Mentions; User	90	157.0	6.0	0.345804	0.370485	0.037553

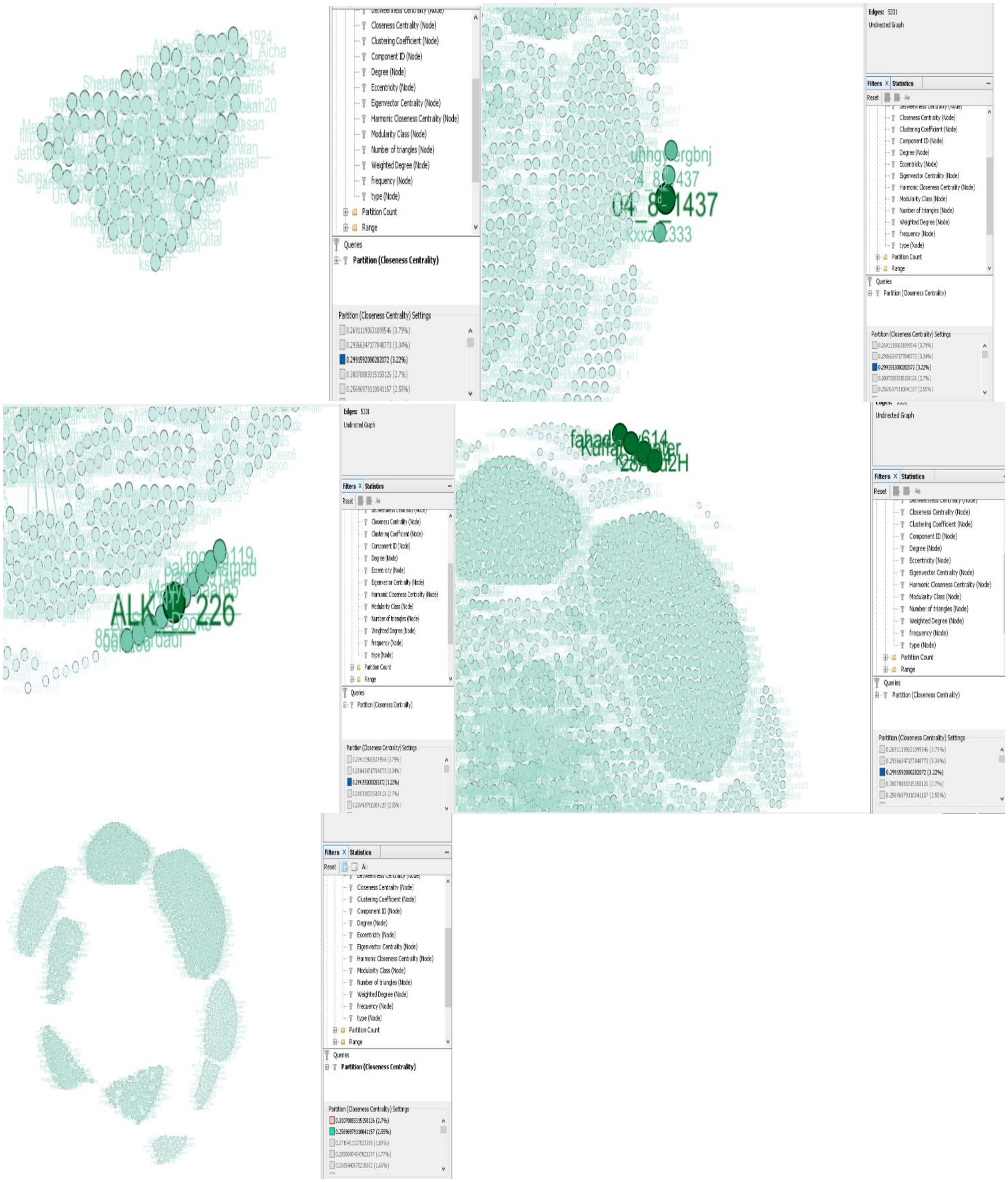
Below is the network Betweenness Centrality:





Below is the network Closeness Centrality:

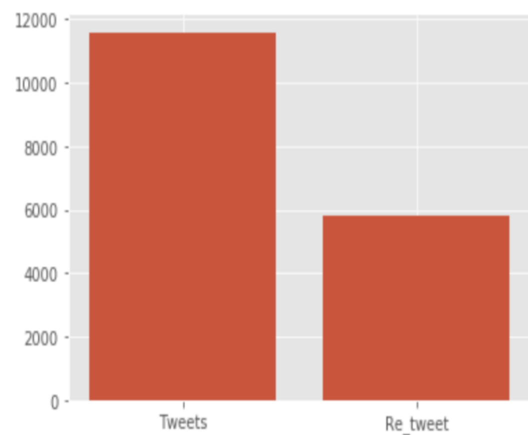



Closeness Centrality (Node)
Clustering Coefficient (Node)
Component ID (Node)
Degree (Node)
Eccentricity (Node)
Eigenvector Centrality (Node)
Harmonic Closeness Centrality (Node)
Modularity Class (Node)
Number of triangles (Node)
Weighted Degree (Node)
frequency (Node)
type (Node)
Partition Count
Range
Queries
Partition (Closeness Centrality)
Partition (Closeness Centrality) Settings
Edges: 5331
Undirected Graph
Filters
Statistics
Reset

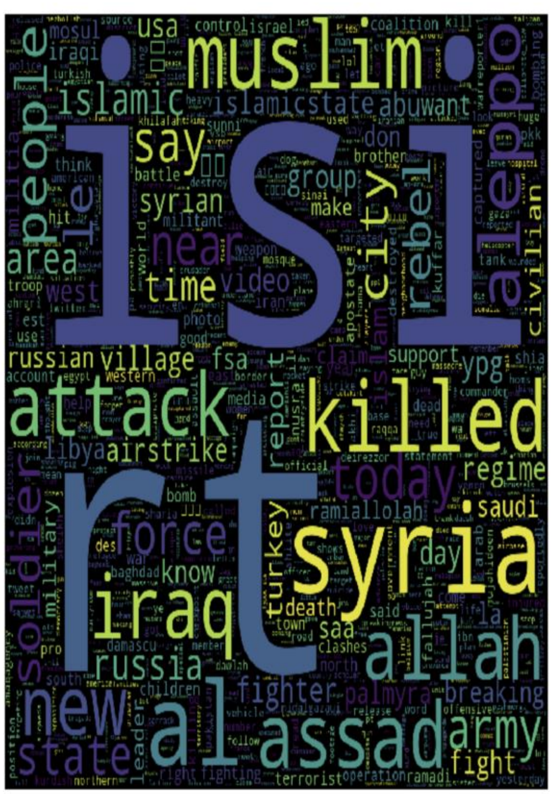
7. Plots and Word Cloud result:

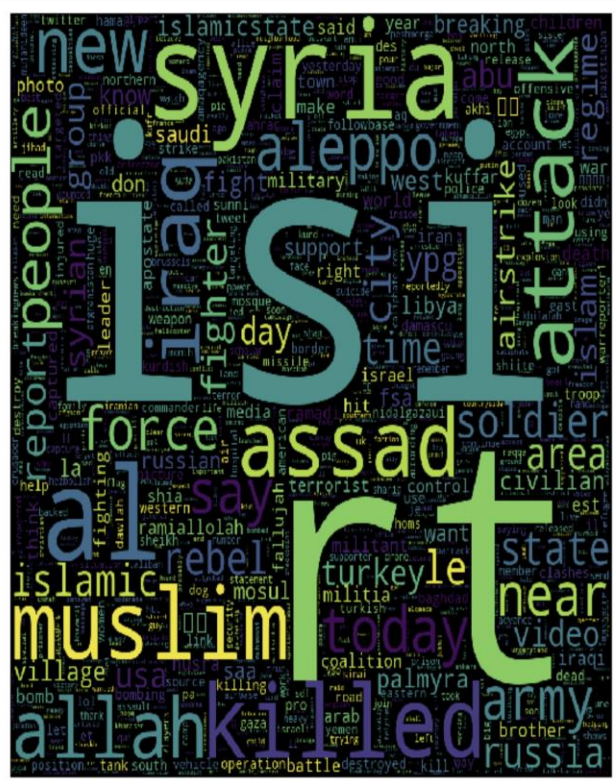
- We can see the below Bar plot about the Actual number of tweets which were done vs the tweets which were re-tweeted by the actors.

```
Out[70]: [Text(0, 0, 'Tweets'), Text(0, 0, 'Re_tweet')]
```



- A word cloud is a popular visualization of words typically associated with Internet keywords and text data. They are most commonly used to highlight popular or trending terms based on frequency of use and prominence. A word cloud is a beautiful, informative image that communicates much in a single glance.[3]





8. Conclusion:

In the recent years, we have seen frequent reports of terrorist attacks all around the world. A good understanding of the terrorist organizations and their social networks is helpful to combat the potential terrorist attacks. Visualization tools are capable to support the analysis of terrorist social networks especially when the networks are large and complex. In this work, we have utilized Fruchterman and Reingold Algorithm to initialize the coordinates of nodes in terrorist social networks and applied the Degree Centrality, Closeness Centrality, Betweenness Centrality, Eigenvector Centrality, Clustering Coefficient, K-core algorithm for visualizing and exploring the global ISIS network interactively. The distance between nodes represents the strength of their associations. Combination of these techniques effectively and efficiently support users to extract to identify the key persons in the terrorist groups and discovering specific patterns of interaction among the terrorists. [20]

9. References and Citations Taken From:

1) https://medium.com/fifth-tribe-stories/isis-has-a-twitter-strategy-and-it-is-terrifying-7cc059ccf51b#.m3zeluykl
2) https://www.kaggle.com/kzaman/how-isis-uses-twitter.
3) https://www.google.com/search?q=what+do+mean+by+word+cloud&rlz=1C1CHBF_enUS899US899&oq=what+do+mean+by+word+cloud&aqs=chrome..69i57j0.6603j0j7&sourceid=chrome&ie=UTF-8
4) https://en.wikipedia.org/wiki/Clustering_coefficient
5) https://en.wikipedia.org/wiki/Degeneracy_(graph_theory)
6) https://faculty.ucr.edu/~hanneman/nettext/C10_Centrality.html
7) https://en.wikipedia.org/wiki/Social_network_analysis
8) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5089772/
9) https://portal.nebih.gov.hu/documents/10182/521653/Algorithms.pdf/6397efcc-6106-4f9e-8057-e7e63858e816
10) https://schneide.blog/tag/fruchterman-reingold/
11) http://cs.brown.edu/people/rtamassi/gdhandbook/chapters/force-directed.pdf
12) https://www.sciencedirect.com/science/article/abs/pii/S2210650215000644
13) https://en.wikipedia.org/wiki/Eigenvector_centrality
14) https://en.wikipedia.org/wiki/Centrality#Harmonic_centrality
15) https://neo4j.com/docs/graph-algorithms/current/labs-algorithms/closeness-centrality/#:~:text=Closeness%20centrality%20is%20a%20way,distances%20to%20all%20other%20nodes.
16) http://matthieu-totet.fr/Koumin/2013/12/16/understand-degree-weighted-degree-betweeness-centrality/#:~:text=Weighted%20Degree%20%3A%20Bigger%20is%20better,the%20weight%20of%20the%20edges.
17) https://www.kaggle.com/jiashenliu/drawing-network-for-pro-isis-tweets
18) https://www.kaggle.com/huwfulcher/social-cluster-analysis
19) https://www.kaggle.com/gruevyhat/militant-extremist-wordcloud
20) Research paper name: Analyzing the Terrorist Social Networks with Visualization Tools
21) Textbook: epdf.pub_social-network-analysis-methods-and-applications-s