Analyzing the Weighted Dark Networks using Scale-Free Network Approach

ABDUL WAHEED MAHESAR[1], AHMAD WAQAS[1,2], NADEEM MEHMOOD[1,3], ASADULLAH SHAH[1], MOHAMED RIDZA WAHIDDIN[1]
[1]Department of Computer Science, Kulliyah of Information & Communication Technology, International Islamic University MALAYSIA
[2]Department of Computer Science, Sukkur Institute of Business Administration, PAKISTAN
[3]Department of Computer Science, University of Karachi, PAKISTAN
abdul.waheed@live.iium.edu.my, nmahmood@uok.edu.pk, ahmad.waqas@live.iium.edu.my, asadullah@iium.edu.my, mridza@iium.edu.my

Abstract: - The task of identifying the main key nodes in the dark (covert) networks is very important for the researchers in the field of dark networks analysis. This analysis leads to locate the major nodes in the network as the functionality can be minimized by disrupting major key nodes in the network. In this paper, we have primarily focused on two basic network analysis metrics, degree and betweenness centrality. Traditionally, both these centrality measures have been applied on the bases of number of links connected with the nodes but without considering link weights. Like many other networks, dark networks also follow scale-free behavior and thus follow the power-law distribution where few nodes have maximum links. This, inhomogeneous structure of network causes the creation of key nodes. In this research, we analyze the behavior of nodes in dark networks based on degree and betweenness centrality measures by using 9/11 terrorist network dataset. We analyzed both these measures with weighted and un-weighted links to prove that weighted networks are much closer to scale-free phenomenon as compared to un-weighted networks.

Key-Words: - Dark networks; power-law distribution; scale-free networks; node degree centrality; betweenness centrality, weighted network analysis.

1 Introduction

A dark network is comprised of actors whose function is to plan and execute any sort of terrorism or criminal activity. These networks evolve slowly from few actors and gradually become the well-organized network with few major nodes and many supporting nodes. Also in such networks, the flow of information between actors occurs when necessary. In complex networks like terrorist networks all the nodes are not equivalent in terms of connection (behavior). Therefore, the removal of nodes having fewer links from the networks has very limited affect whereas as removal of major nodes having maximum links greatly affect the network. The identification of major nodes in dark networks is of primary interest because it may help to minimize the network efficiency. This approach is also used in other types of complex networks such as immunization in networks against epidemics [1] and network tolerance to attacks [2,3].

Links (edges) in between vertices of networks are very fundamental, but all edges/connections are not equal in importance. For example, for two nodes with equal number of links, the node with more powerful links should be more important in network as compared to the one with relatively weak links. Whereas, this phenomenon has not been considered in traditional centralities in case of un-weighted ties among nodes. However, there are few generalizations based on Freeman's centrality measures for weighted networks [4, 5, 6]. The major limitation in all these generalization is its dependence on weight/strength of ties rather than the number of ties, which is the basis for the actual node centrality measures [7]. Terrorist networks are also type of social networks but they significantly differ in terms of their structure as compared to traditional networks [8-16]. Dark networks are also called covert networks because the given information about nodes is often incomplete and insufficient [13,14,15] unlike social networks where we have clearly fixed boundaries.

The monitoring of dark networks is a lengthy and continuous process which takes much longer time to collect secret information. Further, in these types of dynamic networks new vertices and edges in the network are added, which changes the geometry, shape and size of network. This change in the network has a major impact on the analysis approach as it affects the centrality measures of nodes in the connectivity of network [12]. Moreover, the disappearance and appearance of some nodes and ties in a dynamic network majorly affect the centrality measures. Therefore, the modeling and the analysis of dark networks must consider above factors which can make centrality measures of node more robust against minor changes in covert network analysis [17].

The scale-free behavior has been studied in many real world complex networks extensively during the last decade [1,19,20]. We have many examples of complex networks which show the power-law behavior in their formation and growth. Some commonly found networks are Internet, electric power grid, airways, biological, neural and dark networks [33] among many. In scale-free networks, the node degree distribution is not uniform rather it creates inhomogeneous network structure. This proves that few nodes have maximum links compared to many nodes having fewer links in the network. Therefore, if we can locate and disrupt the major nodes in the network with maximum connections along with their weights of connections will greatly reduce the functioning of these networks. The applications of the well-established graph theory and social network analysis are used to identify major nodes in the dark networks. The fig. 1(a) and 1(b) show the distribution of nodes and their links in random and scale-free networks. It is very clear that in case of scale-free networks the distribution of nodes is inhomogeneous whereas in random networks the distribution is fairly equal.

In 1999, Barabasi [28] with the help of WWW map observed that the Internet does not follow random graph connections rather it is scale-free graph and its degree distributions follow power-law form as given in equation (1).

$$P\,(k)\sim\ (k)^{-\gamma} \tag{1}$$

It means, the node degree k and the number of links a node can have, follows the power-law distribution relation where the degree exponent gamma (γ) has been measured as well as confirmed in a number of research studies to be approximately 2.1 [29].

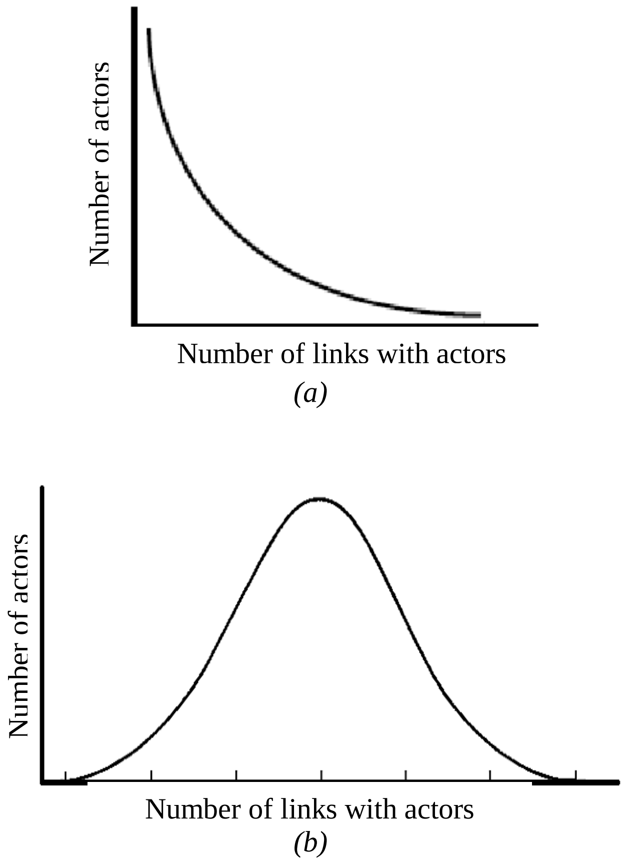


Fig.1 The node degree distribution in (a) scale-free and (b) random networks

For the better and accurate identification of major nodes in dark networks [17] used the value of alpha greater than 1 as compared to [21] in the range of 0 to 2. They showed that high weight of links and nodes play important role in accurate identification of major nodes in covert networks. According to this approach links with high weight play very important role in communication of dark networks. In [30] the authors have focused on the topological information of dark networks based on local and global clustering coefficients. In their method, they focused on the connectivity pattern of nodes from the perspective of close proximity of nodes in the network with important nodes based on the clustering scores of nodes in the network. In [31], the researcher emphasized on the visual analysis of dark networks. According to him, the visual analysis can better explain and predict the major nodes and links in networks with quantitative analysis. The authors in [32] proposed network analysis from the point of view of link prediction when the temporal data of dark networks is available. According to them the valid new missing links can be predicted by using link prediction problem.

In this paper we discuss and highlight the importance of weighted network analysis in case of dark networks by calculating degree and betweenness centrality metrics. Further, we show that the degree and betwenness centrality in dark networks follow power-law distribution and shows the scale-free behavior in the formation and evolution of the dark network. Moreover, we show experimentally that dark networks are very much robust under random attacks due to scale–free topology. This research explicitly focus on the topology of dark networks including other two network analysis metrics as compared to traditional weighted network analysis approaches which solely depends on centrality measures. We have followed the generalized vertex centrality measures given by Opsahl et al. in [21] for calculating the major features and metrics in the network for node centrality. We have used the 9/11 terrorist attack dataset [11] and "r-project" for network analysis.

The rest of this paper is structured as follows. Section 2, discuss and reviews the two metrics degree and betweenness centrality based on [7] and [21] approaches. In Section 3, we analyze the weighted network and importance of including link strengths in the analysis. In section 4, we highlight the features and analysis of 9/11 dark network based on scale-free behavior. Finally, Section 5 concludes the paper with future work.

2 Node Centrality Measurements in Network with Weighted Links

The analysis of node centrality has remained a very important problem in the research community of network analysis [7,22,23]; for the identification of more prominent and central nodes in the network. Because if node is more central as compared to others can have three main advantages. 1) Placement or location can potentially control the traffic/flow among other nodes 2) It can have more direct links 3) From that node we can reach to all other nodes very quickly. Further, these characteristics actually are the basis of Freeman's [7] node centrality measurements namely degree, betweenness and closeness. In this paper, we analyze the behavior of two node centrality measures; degree and betweenness from the perspective of scale-free networks. Moreover, Opsahl et al.[21] has also extended these centrality measurements in case of weighted networks, in which they combined both number and weight of links, by controlling the balance in between these two parameters of major nodes in the networks with the tuning parameter as alpha(α). Here, alpha (α) is the tuning parameter which is a positive quantity and its value can be adjusted according to the context of research setting. Freeman's and Opsahl et al. [21] generalizations give the same results when alpha $(\alpha) = 0.0$. Whereas, on the other hand when alpha $(\alpha) = 1.0$, the obtained values for node centrality are based on purely links weights as given by [4,25].

2.1 The Node Degree of Weighted Networks

The node degree is one of the most basic and simplest indicator of node centrality measures. It is the number of links directly connecting the node with other nodes or simply total number of direct neighbours with node in the network. In network analysis, this metric is mostly used as first step for the analysis due to its simple concept. According to Freeman, node centrality can be mathematically defined as:

$$C_D\,(k) = \, deg_k = \sum_{j}^{N} a_{kj} \tag{2}$$

Where, k is the main node, j represents all other nodes, a is the adjacency matrix where, the entry a_{kj} represent the connectivity if the value is 1, otherwise it is 0, and N is the total number of nodes in the network.

In case of weighted networks, we use the concept of node strength [6], which can be defined as the sum of node direct links weights given in (3)

$$C_D^W(k) = \, strength_k = \sum_{j}^{N} w_{kj} \tag{3}$$

Here, w represents the weighted adjacency matrix. The value of the entry w_{kj} in matrix is > 0, if the node k is connected with j, whereas the value show the weight of link.

Further, in case of un-weighted link between nodes show the presence of relation, and the weight of each link can be taken as 1 therefore, here degree of node and its strength becomes equal e.g, $w_{kj} = a_{kj}$. But in the case of weighted networks the results of both these measures are entirely different and obviously node degree is less preferred measure of strength.

Moreover, as node strength do not consider the number of links with particular node is connected, therefore it gives only rough idea for the nodes actual contribution and involvement in the whole network. By taking simple scenario as given in fig. 2, the three different nodes A,C,and E are equal in strength, whereas we can clearly see that all three are not equally central or important. From these given nodes, the node A has more connections or links with other nodes in the network , therefore we can say node A is more central as compared to C and E.

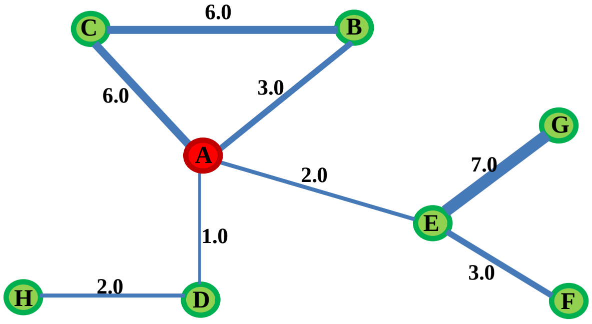


Fig. 2 Example showing the weighted network, with 8 nodes and their links weight

Opsahl et al.[21] has proposed a very realistic approach for degree centrality in weighted networks by considering both number of links and the total weight of links (strength).The reason for this is very much obvious as both these parameters can clearly indicate the involvement of major/focal nodes in overall network [21]. Thus in equation (4)

$$C_D^{W\alpha}(k) = deg_k \times \left(\frac{strength_k}{deg_k}\right)^{\alpha} \tag{4}$$

Here, alpha (α) in this equation works as tuning parameter. If the value of alpha is in the range of 0 and 1, then it will measure both the weights and degree favourably. If the value is more than 1 in that case high valued weights and low degrees are favourably measured [21]. This implies that if we analyze from the perspective of nodes weight then we have to adjust α greatrer than 1, and by increasing the value of α appropriatly we can observe that weighted degree shows scale-free behaviour whereas by decreasing value of α the degree distribution behavoiur changes from scale-free. In this way, we can measure the ranks (importance) of nodes accordigly in the network under observation by using equation (4).

2.2 The Betweenness Centrality in Weighted Networks

A vertex and edge become more important in a network if many shortest paths crosses through that vertex. The betweenness centrality is geodesic-based measure that depends on the finding of shortest paths length and their identification. In simple words, we can say it is a measure to quantify the importance of a vertex in the network on the basis of focal node position on the shortest paths in between the remaining other pairs of nodes. According to Freeman [7] betweenness centrality can be formalised as (5)

$$C_B(k) = \sum_{j}^{N}\sum_{i}^{N}\frac{hji(k)}{hji}\,, i \neq j \tag{5}$$

where h_{ji} is the total number of shortest paths or routes in between two nodes, and $h_{ji}(k)$ is the number of those routes that passes through node k.

Whereas Opsahl et al [21] generalizations of betweenness centrality depends on their generalization of shortest route. The betweenness centrality is formalized according to Opsahl et al. [21] as (6)

$$C_B^{W\alpha}(k) = \sum_{j}^{N}\sum_{i}^{N}\frac{h_{ij}^{W\alpha}(i)}{h_{ij}^{W\alpha}}\,, i \neq j \tag{6}$$

Here, alpha (α) is tuning parameter and when its value is 0 it will calculate the binary shortest distance, whereas in case of 1 it will use Djikstra's algorithm. When the value of alpha is greater than 1, the shortest paths will be based on strongest edges rather than fewest shortest links in between nodes.

2.3 The Shortest Routes in Weighted Network

From the perspective of unweighted network, like in case of distance vector routing the shortest path totally depends on minimum hop count that means less number of intermediary vertices from source to destination is found, and its path or route length is the minimum number of links between these two source and destination. The shortest route between vertices k and j can be defined formally as (7)

$$d(k,j) = min\,(xkh + \ldots..+xhj\,) \tag{7}$$

In this equation h are the nodes that comes in between route of nodes k and j. We can say two direct neighbour can have shortest route of length 1, in case of nodes which are not directly neighbours but both are directly connected with the similar node having shortest path of length 2.

When the links in the network or graph are weighted, then the binary shortest routes are not necessary to be shortest routes, the reason is that the connections and links are different and not all connections can be equally important for the flow of information like in many routing protocols scenario.

For example, different routing protocols find the shortest path based on different weight/metric of the links like bandwidth, delay, speed and congestion control. Therefore, if weight represents strength of links then route composed of high value or strength are shorter as compared to those routes of weaker links. For example, in the network of fig. 3, we have two routes in between two nodes B and C which are connected with different number of intermediary vertices and edges with different weights. The binary shortest route in between these two nodes can be B to C. But as we use the concept of bandwidth in different routing protocols secnario, so the route with high bandwidth will be selected as comapred to direct low bandwidth route. Although the route from (B to D and C) and the route from (B to A to C) has one intermediary route but it can be used as much quicker since they have high value weight. For example, the information can be passed through longer route of strong links more quickly as compared to weaker direct link.

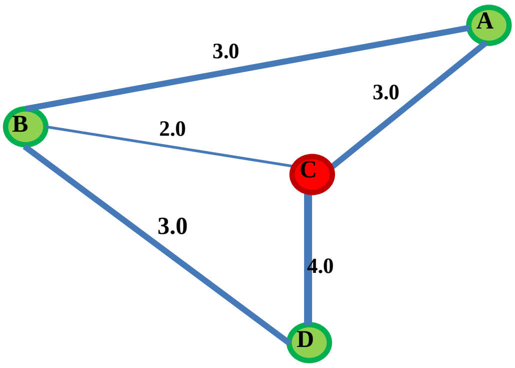


Fig. 3 The weighted network with multiple routes between B & C

Many attempts have been made to find shortest paths in case of weighted networks as given in [5,25,26]. The Dijkstra in [24] proposed the algorithm for finding shortest path, which was used to find shortest path by considering the weight as costs for transferring the messages. Also, well known link state routing protocols are based on this algorithm which are widely used in the field of computer networks today.

3 Network Analyses in Weighted Dark Networks

Actually, the accuracy in measurement of node centrality largely depends on the observed networks data quality. Further, the robustness and reliability under imperfect conditions for node centrality measurement can also affect different characteristics of the network, like network density [23] and its topology [27]. Also, the dark networks are less dense with cellular or clustered topology. Whereas in the networks with high density, high level of accuracy and reliability can be maintained because more links are included or added in the network under observation [23].

Moreover, when the new nodes in observed network can join, the effects will be increasing number of links from the newly coming nodes to already present nodes and new nodes. We can say the addition of number of links in the observed network can be non-linear function of the added nodes like preferential attachment concept.

“When choosing the nodes to which the new node connects, we assume that the probability that a new node will be connected to node i depends on the degree K_j node i,” such that [19]

$$\Pi i = \frac{K_i}{\sum_j K_j} \tag{8}$$

After t time steps, there are

$$N = t + m_0 \text{ nodes} \tag{9}$$

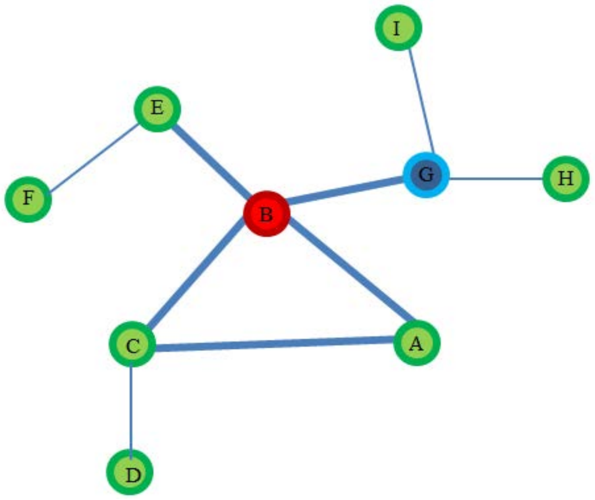
and

$$L = mt + E_0 \text{ edges} \tag{10}$$

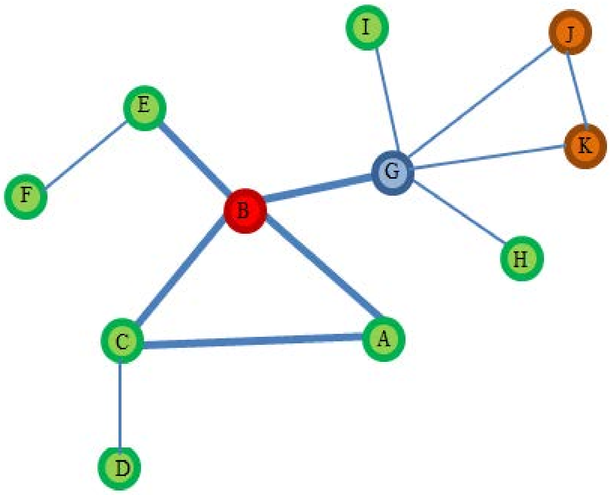
Therefore, in case of dark/terrorist scenario, when sufficient information is obtained or discovered, we can add node and edges to get better and accurate snapshot of the network under observation. Further, when analyzing the terrorist networks, most often the information is incomplete like all of its constituent nodes, the relations and links in between those nodes and actual structure of network. But we can follow the trend from previous or initial data about the growing nature of the network like preferential attachement rule. Therefore, we can’t rely hundred percent on given data because true network can be different from observed network.

Now, fig. 4 shows two scenario about the different analyzed states in a weighted covert network. Here, the bold links shows the connections with double impact of strength as compared to thinner links. In first scenario figure 4(a) the node B has greater degree distribution or number of links therefore B is the most important or central vertex from the point of view of weighted and unweighted measures of centralities. Whereas node G which is second largest node degree in the network but it has weak links or ties. Therefore, node B is more important with highest number of links as well as strength of links. Now in figure 4(b) if we add two more nodes to G and if we ignore the weights in the network, we can see that node G becomes more central than node B previously mentioned. But, if we don't ignore the links weights, and we try to find out the stronger instead of highest links by controlling both centrality measures by adjusting the tuning parameter alpha greater than 1, then node B will remain the most focal/central node equivalent with the observation of result in fig 4(a). Therefore, when new nodes or links with high or stronger weigths are added this can affect the centrality otherwise the analysis in case of weak links can not be affected.

The same observation is verified and tested in weighted dark network with real dataset, that was actually used and created by Krebs [11] based on 9/11 incident in which we have 62 actors/nodes and 153 links/ties with high, medium and low strengths. We assume that it is a true network and from that network we will drive a network which can be an observed network. For obtaining the observed network, we extract the samples from the original or true network by randomly removing 5%,10% and 15% low-strength links from the given network. Here, our main focus is to see the generalized centrality of node rankings in actual network with obtained rankings by changing the links percentage in all cases by taking different values of alpha. More specifically, our aim is to see the correlation in between true node centrality and observed networks by changing alpha when it is 0.0, 0<alpha<1 , alpha=1 and finally when alpha>1.



(a)



(b)

Fig 4(a & b) Example showing effects of changes in the network with node centrality.

We try to analyze the nodes linkages behaviour with weighted and unweighted links and specifically show that weighted dark networks follow scale-free behaiour. We have used Spearman's Rank correlation for finding correlation coefficient (ρ) inbetween true network and observed networks by randomly removed links. The experiments have been performed three times and values are averaged as shown in table 1. Also, we have used 0.5 value interval from 0.0 to 3.0, for each value of alpha. The computed correlation coefficient will show the difference in between true/actual network and the network by taking 5%,10% and 15% missing links in observed network.

From table 1 we can compare the similarity in between actual and three observed networks when we remove 5%,10% and 15% links from the actual network. This is done by observing the ranking of node centrality when the values of α changed from 0.0 to 3.0. The values in case of 5%,10% links removed from network (Net)1 and (Net)2 are higher than 15% (Net)3 links removed which clearly shows that (Net)1 is closer to original network. Also, we can see that when we fully ignore the links weights the correlation is very low for both degree and betweenness centralities. But in case of higher values of alpha we get greater values of correlation as we are including the weights of the links in the analysis. Moreover, as found from literature that the scale-free networks are robust under random links failure and are very fragile or vulnerable under intentional or targeted attacks, so we can see that the original network sustain that property more closely of scale-free behavior when it is analyzed from the weighted links perspective while in case of un-weighted links the above network behavior is not closer to scale-free.

Table 1 The effect of random links removal from the original network on degree and betweenness centralities using Spearman's rank correlation as compared to true network.

Tuning Parameter	**5 % connections removal (Network1)**	
α	Degree	Betweenness
0.0	0.46	0.55
0.5	0.41	0.49
1.0	0.60	0.55
1.5	0.67	0.59
2.0	0.76	0.60
2.5	0.71	0.55
3.0	0.77	0.70
	10 % connections removal (Network2)	
0.0	0.24	0.55
0.5	0.35	0.36
1.0	0.56	0.57
1.5	0.57	0.44
2.0	0.64	0.58
2.5	0.68	0.55
3.0	0.69	0.69
	15% connections removal (Network3)	
0.0	0.28	0.49
0.5	0.32	0.58
1.0	0.33	0.50
1.5	0.38	0.59
2.0	0.45	0.50
2.5	0.46	0.49
3.0	0.57	0.59

4 A look at 9/11 Network: (Features, Metrics, and Weights)

This network shows many features of dark/criminal networks. There are 62 nodes/vertices and 153 connections [11]. Although the density of this sample network is not high and there exists 8% possible conections/links. The reason behind that low density is quite obvious as dark networks try to focus on secrecy more as compared to efficiency [11]. This is the reason for the behaviour of dark networks which shows the scale-free nature as there are few nodes or actors with high number of links, and therefore random link failure does not effect on their functionality. By identifying the core nodes in the network and removing them from network speadily collapes the functionality of dark networks. In this scenario, the network has 4 clusters/groups and 19 major actors which are tightly linked with each other. Here, these 4 clusters/groups belongs to the persons who were present in the hijacking of US planes. This network has average degree distribution of 4.9 that means we can reach to approximately five actors from each one actor of the network. Also, the metric diameter of the network is five. This diameter represents the degree separation is five in between any two actors.

As for as weight is concerned, in this paper we have used Krebs network dataset that is based on weights [11],[17].The links between the individuals in this paper represent the kind of interaction that happened previously. The weights of links are based on 1 to 3 scale, and it can be many factors like relationship between links, association between links, trust between links and many more. The major in all factors is how much time two actors in the network has spent together [11]. Here, we can say strongest links are based on how much time two nodes have spent together for example studied in the same institute. For average or modest links weight the participation in same meeting or training is considered. And for weaker links the weight represents just the occasional meeting and rare interaction[11]. For calculating the values of centralities we have used open source software package the r- Project, which is an open source software package. We have used data from the Krebs [11] which is cited in many studies based on 9/11 incident and it is publicly available.

4.1 Outcomes and Analysis

The selection for the values of alpha depends on many factors like research context, the available network data, and features of datasets. For example, the variation in links strength. According to Opsahl et al. [21] by setting the values of alpha as 0.0 the result will show the unweighted network and by selecting alpha as 1.0 will give centralities values only on the basis of nodes links weights. On the other hand, for values of alpha greater than 1.0, the node centralities will be calculated by considering the degree and total weight/strength of its links.

In case of our analysis, we have set the value of alpha to 2.0. The major reason for selecting the alpha equal to 2.0 is as given by [11] that the communication happens between those nodes which are based on high level of trust. It means that most important communication occurs inbetween those nodes which are strongly connected in the network with high strengths. Further, it is obvious that if a node has many connections/links then it will be very much vulnerable to discovery. This is the reason for selecting value of alpha >1, so that to focus on stronger links of nodes as compared to many links of nodes. Similarly, for selecting shortest paths a path with stronger links as compared to fewer links to be analyzed.

Table2 The degree centrality of the Top fifteen nodes in the network when alpha (α) = 0.0 and 2.0.

Number of Node	**The effect on degree centrality when (alpha=0.0)**		**The effect on degree centrality when (alpha=2.0)**	
	Name of Node	*Degree*	*Number of Node*	*Degree*
1	Mohamed Atta *	0.361	1*	0.966
2	Marwan Al-Shehhi **	0.295	6 ****	0.867
3	Hani Hanjour ***	0.213	04***	0.858
4	Nawaf Alhazmi ***	0.180	02 **	0.820
5	Essid Sami Ben Khemais	0.180	Zakaryia Essabar	0.738
6	Ziad Samir Jarrah ****	0.164	03***	0.726
7	Ramzi Omar	0.164	14	0.677
8	Abdulaziz Alomari *	0.148	07	0.656
9	Satam M. A. Al Suqami *	0.131	Saeed Alghamidi****	0.615
10	Salem Alhazmi ***	0.131	Hamza Alghamdi **	0.527
11	Fayez Rahq Banihammad **	0.131	Ahmed Alnami	0.443
12	Habib Zacarias Moussaoui	0.131	Ahmed Ibrahim	0.410
13	Djamal Benghal	0.131	09	0.402
14	Said Bahaji	0.115	Khalid	0.393
15	Hamza Alghamdi **	0.115	Lotifi Raissi	0.328

*American Airline Flight 11 (Crashed into WTC 1) **United Airline Flight 175 (Crashed into WTC 2)
American Airline Flight 77 (Crashed into the Pentagon) * United Airline Flight 93 (Crashed in Pennsylvania)

Table3 The betweenness centrality of Top fifteen nodes in the network when α= 0.0 and 2.0.

Number of Node	**The effect on betweenness centrality when (alpha=0.0)**		**The effect on betweenness centrality when (alpha=2.0)**	
	Name of Node	*Betweenness*	*Number of Node*	*Betweenness*
1	Mohamed Atta *	0.561	1	0.560
2	Essid Sami Ben Khemais	0.243	2	0.256
3	Habib Zacarias Moussaoui	0.233	12	0.251
4	Hani Hanjour ***	0.154	AhmedIbrahimHaznaw****	0.248
5	Nawaf Alhazmi ***	0.136	4	0.240
6	Marwan Al-Shehhi **	0.116	3	0.230
7	Djamal Benghal	0.104	Saeed Alghamadi ****	0.167
8	Satam M. A. Al Suqami *	0.048	5	0.141
9	Ramzi Omar	0.046	Hamza Alghamdi **	0.108
10	Abu Qatada	0.039	7	0.102
11	Tarek Maaroufi	0.038	Nabil al-Marabh	0.085
12	Ziad Samir Jarrah ****	0.037	9	0.084
13	Mamoun Darkazanli	0.033	8	0.063
14	Imad Eddin Barakat Yarkas	0.033	10	0.036
15	Fayez Rashid Banihammad **	0.026	Khalid Almindhar ***	0.036

*American Airline Flight 11 (Crashed into WTC 1) **United Airline Flight 175 (Crashed into WTC 2)
American Airline Flight 77 (Crashed into the Pentagon) * United Airline Flight 93 (Crashed in Pennsylvania)

Fig. 5 depicts the network topology of 9/11 network that shows the behaviour of nodes and links where few nodes have many links and majority have very few. Also the weight of these links varies significantly with low to high.

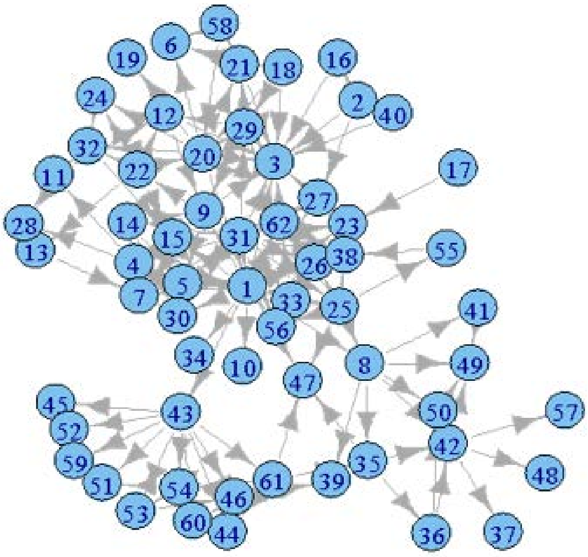


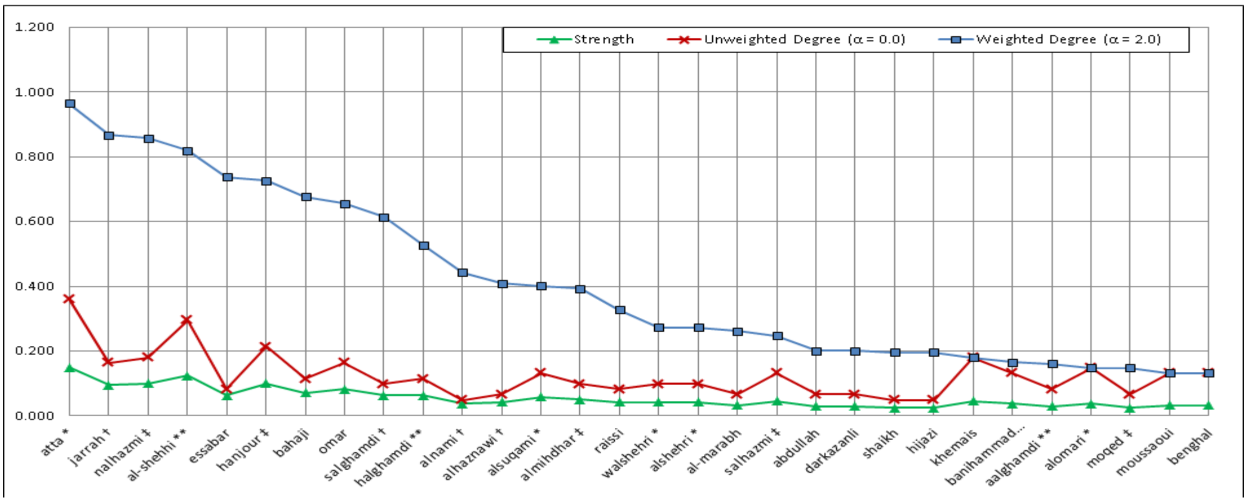
Fig.5 Network visualization in r-project of 9/11 network.

By usnig the generalized centrality measures as given by Opsahl et al.[21] and selecting alpha=2.0, we have calculated the degree and betweenness centrality of 9/11 network. For analysis purposes, tables 2 and 3 shows the top fifteen nodes based on degree and betweenness centralities. We have listed the top 15 nodes from highest to lowest values in un-weighted and weighted network centralties. The above table shows the effect on ranks of actors (with their names and degree centrality) when we are including the links only (α=0.0), and when we are including weight of the links in analysis (α=2.0) neglecting the number of links. The columns in the table represent the node numbers, names and their degree centralities rank wise. The results shows clearly that for example node number 6 named (Zaid bin Jarrah) is on rank 6 when α=0.0 and he becomes the second number in rank when α=2.0.

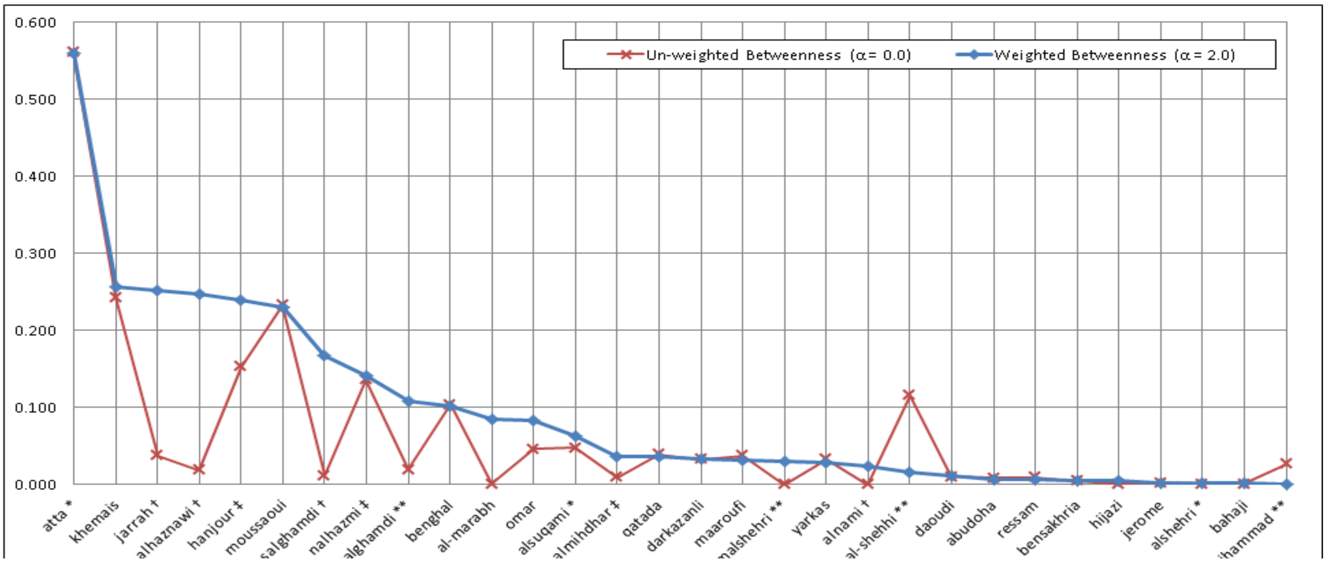
Moreover, few more new actors appear in the list of ranks due to inclusion of weight of links for example Zakariya Essabar who is above fifteen numbers when we have analysed form links point of view but he is on fifth number when we have analysed by including weight only.

We can see that in weighted and un-weighted analysis the node ID 1 ‘Atta’ is the most important and central node in network from degree and betweenness perspectives. If we look at top 4 nodes in table 2 these four nodes are most central according to degree and each node belong to different group of hijackers. Also, from the betweenness perspective, in table 3 Essid and Habib are on top, and their function is to connect the nodes with other associate nodes in the network. Further, the node 6 Zaid Samir is more important when analysed with weighted centrality, representing that he is linked through stronger links/connections with other important nodes in the network. It means that he is connected with stronger links and the distance between nodes from Zaid to others are relatively strong links.

Moreover, weighted network analysis shows trend towards scale-free behaviour as shown in the fig. 6. We have shown the top 30 nodes from the perspective of values of alpha=2.0. The x-axis represents the nodes and y-axis represents the centralities of network in these graphs (a) and (b). For better comparison we have plotted the un-weighted centrality values as well. Here, we can see clearly that the nodes are decreasing with the decreasing centrality and the decreasing curves show more closer trend to the power-law behaviour in case of weighted network.



(a) Degree Centrality



(b) Betweenness Centrality

Fig.6 Graphs (a) and (b) showing the weighted and un-weighted dark networks comparison. Both graphs are plotted by top thirty actors sorted by weighted centralities.The grpah shows satisfactorily results when the weight of links is increased it is very close to scale-free behavior as compare to unweighted network analysis.

As the nodes and links in this network are 62 and 153 which is not that large, therefore we can not find the value of gamma by plotting dataset on log-log scale. But from the available dataset it is clear that there are very few nodes with very high centrality and many are with less.

5 Conclusions

In case of dark networks analysis like homeland security, Drug trafficking gangs, human trafficking groups and terrorist organizations etc. the correct understanding of data gathered, intelligence reports and proper analysis is very important for accurate prediction of networks activity. Now a dyas national security has become very important area of reaserch mainly after the terrorism activities around the globe. Therefore, counterterrorist organizations are focusing on the data gathered from such types of dark networks from different perspectives to prevent any disasterous event in future. This is the main reason of popularity of social network analysis as very active area of research.

In our paper, we have analyzed the dark network by using social network analysis metrics with scale-free phenomenon. For major nodes identification in the network and the robustness sustainability, scale-free network analysis approach has been adopted. The main reason of introducing this approach is an effecient and accurate identification of major nodes in the network. The proposed approach has been applied as a case study of 9/11 terrorist network dataset. The outcome of analysis shows satisfactorily results as the above network sustain its robustness by including weight of links in the analysis. Also, the two network analysis metrics when analyzed from weighted liks point of view identify accurately major nodes in the network. Further, the observed behaviour of the metrics degree and betweenness shows that it is true for dark networks to have scale-free behaviour in their formation as very few nodes are very important and many nodes functions as supporting or they are at periphery of the network as compared to main or central nodes. By using "r-project" for analysis and plotting the graphs we have obtained the satisfactorily results which shows that weighted networks are very close to scale–free nature as compared to un-weighted dark networks. Moreover, the correlation in table 1 shows that this network is very much robust under random link failure or removal.

6 Future Work and Improvements

The limitation of this analysis is the lack of different dark networks dataset and their availability. In particular case of Kerbs' dataset we have obtained encouraging outcomes which shows the approach is positive. The limitation of huge dataset in terms of nodes and links make it difficult to find power-law exponent which can accurately predict the topology in terms of power-law distribution. On the other hands our results shows satisfactorily the trend of decreasing curve which can be predicted as power-law. We believe that if this network grows in this pattern then it may follow rich get richer phenomenon. There are many other complex network analysis metrics that can be verified for the scale-free behaviour in dark networks like clossenes, distance, clustering, homophile and these may constitute future work with few other dataset of dark networks.

References:

[1] Pastor-Satorras, Romualdo, and Alessandro Vespignani. Epidemic spreading in scale-free networks, *Physical review letters* 86.14 (2001): 3200.

[2] J. M. Clark, D. A. Holton. *A first look at graph theory*, Singapore : World Scientific, 1991

[3] Albert, Réka, Hawoong Jeong, and Albert-László Barabási. Internet: Diameter of the world-wide web, *Nature* 401.6749,1999, pp. 130-131.

[4] U. Brandes, A faster algorithm for betweenness centrality, *Journal of Mathematical Sociology*, vol. 25, 2001, pp. 163–177.

[5] M. E. J. Newman. Analysis of weighted networks, *Physical Review E* 70.5 (2004): 056131.

[6] A. Barrat, M. Barthlmy, R. Pastor-Satorras, and A. Vespignani, The architecture of complex weighted networks, *Proceedings of the National Academy of Sciences*, vol. 101 (11), 2004, pp. 3747-3752.

[7] L. C. Freeman, Centrality in social networks: conceptual clarification, *Social Networks*, vol. 1, 1978, pp. 215–239.

[8] M. Sageman, *Understanding Terror Networks*, University of Pennsylvani Press, Philadelphia, 2004.

[9] R. Lindelauf, P. Borm, and H. Hamers, Understanding terrorist network topologies and their resilience against disruption" Counterterrorism and Open Source Intelligence, 2011, pp. 61-72.

[10] C. T. Butts, The complexity of social networks: Theoretical and empirical findings, *Social Networks*, vol. 23 (1), 2001, pp. 31-71.

[11] V. E. Krebs, Uncloaking terrorist networks, *First Monday*, vol. 7, pp.4–11, 2002.

[12] C. T. Butts, Network inference, error, and informant (in) accuracy: a Bayesian approach, *social networks*, Vol. 25 No. 2, 2003, pp. 103-140.

[13] K. Carley, "Vulnerabilities in large covert networks," in *Proceedings of the NAACSOS 2004 Conference*, Pittsburgh, PA, 2004.

[14] K. Carley, J.-S. Lee, and D. Krackhardt, Destabilizing networks, *Connections*, vol. 24 No. 3, 2001, pp. 31-34.

[15] K. Carley, M. Dombroski, M. Tsvetovat, J. Reminga, and N. Kamneva, Destabilizing dynamic covert networks, in *Proceedings of the 8th International Command and Control Research and Technology*, National Defense War College, Washington DC, 2003.

[16] Philip V. Fellman and Roxana Wright, "Modeling Terrorist Networks: Complex Systems at the Mid-Range", Joint Complexity Conference, London School of Economics, September, 2003.

[17] Memon, Bisharat Rasool. Identifying important nodes in weighted covert networks using generalized centrality measures." Intelligence and Security Informatics Conference (EISIC), 2012 European. IEEE, 2012.

[18] S. H. Strogatz, Exploring complex networks, *Nature*, vol. 410, 2001, pp.268-276.

[19] R. Albert and A-L. Barabási, Statistical mechanics of complex networks, *Review of Modern Physics*, vol. 74, 2002, pp. 47-91.

[20] X. F. Wang, Complex networks: topology, dynamics and synchronization, *Int. J. Bifurcation & Chaos*, vol. 12, No. 5, 2002, pp. 885-916.

[21] T. Opsahl, F. Agneessens, and J. Skvoretz, "Node centrality in weighted networks: Generalizing degree and shortest paths," Social Networks, vol. 32, 2010, pp. 245-251.

[22] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994.

[23] S. P. Borgatti, K. Carley, and D. Krackhardt, On the robustness of centrality measures under conditions of imperfect data, *Social Networks*, vol. 28, pp. 124-136, 2006.

[24] E. W. Dijkstra, A note on two problems in connexion with graphs, *Numerische Mathematik*, vol. 1, 1959, pp. 269-271.

[25] M. E. J. Newman, Scientific collaboration networks. II. shortest paths, weighted networks, and centrality, *Physical Review*, vol. E 64, 2001, p.016132.

[26] L. C. Freeman, S. P. Borgatti, and D. R. White, Centrality in valued graphs: a measure of betweenness based on network flow, *Social Networks*, vol. 13 No. 2, 1991, pp. 141–154.

[27] T. L. Frantz and K. Carley, "Relating network topology to the robustness of centrality measures," No. CMU-ISRI-05-117. CARNEGIE-MELLON UNIV PITTSBURGH PA SCHOOL OF COMPUTER SCIENCE, 2005.

[28] A-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science*, vol. 286, pp. 509-512, 1999.

[29] Goh, Kwang-Il, Byungnam Kahng, and Doochul Kim. *Evolution of the Internet topology and traffic dynamics of data packets*. Complex Dynamics in Communication Networks. Springer Berlin Heidelberg, 2005. 235-250.

[30] Xu, Jennifer, et al. "On the topology of the dark web of terrorist groups."*Intelligence and Security Informatics*. Springer Berlin Heidelberg, 2006. 367-376.

[31] Everton, Sean S. "Tracking, Destabilizing and Disrupting Dark Networks with Social Networks Analysis." (2008).

[32] Grady, Daniel, Christian Thiemann, and Dirk Brockmann. "Robust classification of salient links in complex networks." *Nature communications* 3 (2012): 864.

[33] A.W. Mahesar, A. Messikh, A. Shah, M.R. Wahiddin "Node status detection and information diffusion in router network using Scale-free network" *Proceedings of the Tenth International Network Conference (INC 2014) Plymouth*, UK, July 8-10, 2014.

Appendix:

The Table given below shows the IDs of actors/members and their respective names in 9/11 dark network [11].

NodeID	Name	NodeID	Name	NodeID	Name
1	Ziad Samir Jarrah****	26	Ramzi Omar	51	Kamel Daoudi
2	Mohamed Atta*	27	Said Bahaji	52	Lased Ben Heni
3	RayedMohammed Abdullah	28	Lotfi Raissi	53	Madjid Sahoune
4	Hani Hanjour***	29	Raed Hijazi	54	Mehdi Khammoun
5	Satam M. A. Al Suqami*	30	Salem Alhazmi***	55	Mohamed Bensakhria
6	Wail M. Alshehri*	31	Shaykh Saiid	56	Mohammed Belfas
7	Abdussattar Shaikh	32	Marwan Al-Shehhi**	57	Mounir El Motassadeq
8	Fayez Rashid Banihammad**	33	Saeed Alghamdi****	58	Nizar Trabelsi
9	Habib Zacarias Moussaoui	34	Zakariya Essabar	59	Osama Awadallah
10	Abdulaziz Alomari*	35	Abdelghani Mzoudi	60	Samir Kishk
11	Ahmed Khalil Ibrahim Samir	36	Abu Qatada	61	Seifallah ben Hassine
12	Nabil al-Marabh	37	Abu Walid	62	Tarek Maaroufi
13	Ahmed Alghamdi**	38	Abu Zubeida		
14	Mohand Alshehri**	39	Agus Budiman		
15	Waleed M. Alshehri*	40	Ahmed Ressam		
16	Ahmed Ibrahim A. Al Haz****	41	Bandar Alhazmi		
17	Faisal Al Salmi	42	David Courtaillier		
18	Mamduh Mahmud Salim	43	Djamal Benghal		
19	Majed Moqed***	44	Essid Sami Ben Khemais		
20	Mohamed Abdi	45	Essoussi Laaroussi		
21	Nawaf Alhazmi***	46	Fahid al Shakri		
22	Khalid Almihdhar***	47	Haydar Abu Doha		
23	Hamza Alghamdi**	48	Imad Eddin Barakat		
24	Mamoun Darkazanli	49	Jean-Marc Grandvisir		
25	Ahmed Alnami****	50	Jerome Courtaillier		

*American Airline Flight 11 (Crashed into WTC 1) *** American Airline Flight 77 (Crashed into the Pentagon)
** United Airline Flight 175 (Crashed into WTC 2) **** United Airline Flight 93 (Crashed in Pennsylvania)