Investigative Data Mining Toolkit: A Software Prototype for Visualizing, Analyzing and Destabilizing Terrorist Networks

Nasrullah Memon and Henrik Legind Larsen
Software Intelligence Security Research Center
Department of Computer Science and Engineering
Aalborg University, Niels Bohrs Vej 8
6700 Esbjerg, Denmark
nasrullah@cs.aaue.dk, legind@cs.aaue.dk

ABSTRACT

Knowledge about the structure and organization of terrorist networks is important for both terrorism investigation and the development of effective strategies to prevent terrorists' attacks. However, except for network visualization, terrorist network analysis remains primarily a manual process. Existing tools do not provide advanced structural analysis techniques that allow extraction of network knowledge from large volumes of criminal-justice data. It is a well known fact that terrorist activities consist of dispersed organizations (like non-hierarchical organizations), small groups, and individuals who communicate, coordinate and conduct their campaign in a network-like manner. There is a pressing need to automatically collect data of terrorist networks, analyze such networks to find hidden relations and groups, prune datasets to locate regions of interest, find key players, characterize the structure, trace point of vulnerability, and detect efficiency of the network. To meet this challenge, we designed and developed a knowledgebase for storing and manipulating data collected from various authenticated websites. This paper presents framework of investigative data mining toolkit, our recently introduced techniques and algorithms (which are implemented in the investigative data mining toolkit) could be useful for law enforcement agencies that need to analyze terrorist networks and prioritize their targets. Applying recently introduced algorithms for constructing hidden hierarchy of non-hierarchical terrorist networks, we present case studies of the terrorist attacks that occurred in past, in order to construct command structure of the networks.

KEYWORDS

Visualization, Analysis and Destabilizing Terrorist Networks, Position Role Index, Dependence Centrality, Hierarchy of Non-Hierarchical Networks

1.0 INTRODUCTION

The threats facing society today require new methods for modeling and analysis. In fact, civil security decision makers, analysts and field operators fighting terrorism and organized crime across the European Union all need front-line integrated technologies to support their cooperative work. Our opponents are no longer organized in hierarchical structures, but instead consist of individuals and groups that are loosely organized in "dark networks". Instead of large-scale military attacks, they stage attacks or set bombs against unprotected civilians, or seek to influence crowds of legitimate demonstrators so that critical riot situations occur.

In order to construct decision support systems that take account of these new factors, new, more powerful methods and techniques from several technological domains need to be brought together and integrated. Experience shows that the networks can be unwound and analyzed after the events. Although it provides the necessary evidence for bringing criminals to justice, it is then too late to prevent loss of life and

Memon, N.; Larsen, H.L. (2006) Investigative Data Mining Toolkit: A Software Prototype for Visualizing, Analyzing and Destabilizing Terrorist Networks. In *Visualising Network Information* (pp. 14-1 – 14-24). Meeting Proceedings RTO-MP-IST-063, Paper 14. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

material damage.

Mathematical methods used in our research on Investigative Data Mining [1] [2] are clearly relevant to law enforcement intelligence work and may provide tools to discover terrorist networks in their planning phase and thereby prevent terrorist acts and other large-scale crimes from being carried out. Relevant patterns to investigate include connections between actors (meetings, messages), activities of the involved actors (specialized training, purchasing of equipment) and information gathering (time tables, visiting sites).

Investigative Data Mining (IDM) offers the ability to firstly map a covert cell, and to secondly measure the specific structural and interactional criteria of such a cell. This framework aims to connect the dots between individuals and “map and measure complex, covert, human groups and organisations”. The method focuses on uncovering the patterning of people’s interaction, and correctly interpreting these networks assists “in predicting behaviour and decision-making within the network”.

IDM also endows the analyst the ability to measure the level of covertness and efficiency of the cell as a whole, and also the level of activity, ability to access others, and the level of control over a network each individual possesses. The measurement of these criteria allows specific counter-terrorism applications to be drawn, and assists in the assessment of the most effective methods of disrupting and neutralising a terrorist cell. In short IDM “provides a useful way of structuring knowledge and framing further research. Ideally it can also enhance an analyst’s predictive capability”.

IDM borrowed social network analysis (SNA) and graph theory techniques for connecting the dots; our goal is to propose mathematical methods for destabilizing terrorist networks after linking the dots between them.

In investigative data mining, a number of variations exist in the literature. One is known as link analysis (see for example [5] [6]). Link analysis research uses search and probabilistic approaches to find structural characteristic in the network such as hubs, gatekeepers, pulse-takers [7], or identifying potential relationships for relational data mining. Link analysis alone is insufficient as it looks at one side of the coin and ignores complex nonlinear relationships that may exist between the attributes. Another approach depends purely on visualization, such as NetMap [8]. Unfortunately, these tools that depend on visualization alone - despite being useful to provide some insight - they are insufficient and rely on the user to carry out many tedious and time consuming tasks, many of which could be automated.

In addition to the previous discussion, most of the work on link analysis or network visualization ignores the construction of hidden hierarchy of coverts networks. Uncovering a relationship among or within attributes (connecting the dots) is an important step, but in many domains it is more important to understand how this relationship evolved. Hence, understanding network dynamics and evolution is needed to complete the picture. Once we understand the dynamics and evolution of these relationships and construct the hidden hierarchy, we can search for ways to disconnect the dots when and if needed.

The three innovative points of our research are [2]:

1. The use of new measure Position Role Index (PRI) on the pattern of efficiency introduced by Vito Latora and Massimo Marchiori [27]. This measure identifies key players (gatekeepers/ leaders) and followers in the network.
2. The use of another measure known as Dependence Centrality (DC) which discovers who is depending on whom in a network.
3. Estimate possible hierarchical structure of a complex network by applying degree centrality and Eigenvector centrality from social network analysis (SNA) literature and combining it with new measure dependence centrality.

This paper presents some case studies of the terrorist events occurred in the past, using software prototype that we have developed. The structure of the rest of this paper is as follows. In Section 2, a brief introduction about terrorist network analysis is presented. Section 3 presents a concise review to point the reader to several key papers in the literature; whereas Section 4 discusses the models we have used for destabilizing terrorist networks. Section 5 discusses an overview of the iMiner software prototype system; whereas four case studies are then presented in Section 6. Conclusions are then drawn in Section 7.

2.0 TERRORIST NETWORK ANALYSIS

The threat from modern terrorism manifests itself worldwide in locally and internationally operating network structures. Before focusing on the development and composition of these networks, we will initially attempt to answer the question: what exactly is a terrorist network?

Persons involved in support, preparation or commission of terrorist attacks almost never operate alone, but as members of - sometimes overlapping - network structures. Within these networks they co-operate with individual members or small groups of members (operational cells). A modern terrorist *network* differs from other terrorist groups and organizations in that it lacks a formal (hierarchical) structure, and has an informal, flexible membership and fluctuating leadership. It is incorrect, however, to conclude that such a network possesses no structure whatsoever. There is always a pattern of connections between individuals who communicate with one another with a view to achieving a common goal. In some cases these communication lines converge in one or more core groups, which thus play a coordinating and controlling role. In other cases there are random communication patterns between all members while the network functions practically without any leadership or central control. It is also possible for several groups to be active within one network.

The flexible and informal character of such a network makes it easy for individual members to establish temporary ad-hoc contacts, in addition to more permanent relations. It also leaves room for personal initiative. The relations within a network are constantly changing in character and duration. In most cases we can distinguish a core group surrounded by a diffuse network of individuals, with central control usually restricted to a minimum. Personal ties between members bind the network together. These relationships are usually based on a shared political-religious ideology, mutual trust, family or friendship ties, shared origin and/or shared experiences in training camps or jihad areas. The notion of a common *enemy* also stimulates bonding among network members.

The above characteristics lead to the following definition:

A terrorist network is a fluid, dynamic, vaguely delineated structure comprising a number of interrelated persons who are linked both individually and on an aggregate level (cells / groups). They have at least a temporary common interest, i.e. the pursuit of a jihadism-related goal (including terrorism).

Persons within such a network are referred to as members. A member is a person who contributes actively and consciously to the realization of the aforementioned goal within the bounds of the network.

This definition is in line with the definition of criminal networks used in Criminology, which does not refer to permanent structures, but to temporary, flexible co-operative structures between individuals, based on kinship, friendship, business opportunism, coincidence, necessity, temptation and force, or to the fact that members are colleagues, neighbours or fellow convicts. This co-operation gradually evolves into certain customs and traditions which lead to ‘habituation, mutual interdependence and trust, and hierarchical relations [8]. This assessment of fluid and dynamic criminal networks was described in an extensive study into organized crime [9].

2.1 Centrality measures for Analyzing Terrorist Networks

Centrality is one of the most important and widely used measures for analyzing social networks. Nearly all empirical studies try to identify the most important actors (also known as vertices / nodes) within the network. Four measures of centrality are commonly used in network analysis: degree, closeness, betweenness, and eigenvector centrality. The first three were described in modern form by Freeman [10] while the last was proposed by Bonacich [11]. Let us begin with degree.

Degree centrality measures how active a particular node is. It is defined as the number of direct links a node k has:

$$C_D(k) = \sum_{i=1}^{n} a(i,k), \qquad (1)$$

where *n* is the total number of nodes in a network, and *a(i ,k)* is a binary variable indicating whether a link exists between *i*th and *k*th nodes. A network member with a high degree could be the leader or "hub" in a network.

Closeness centrality [10] is the sum of the length of geodesics (shortest paths between two nodes) between a particular node k and all other nodes in a network. It actually measures how far away one node is from other nodes and is sometimes called *farness*:

$$C_c(k) = \sum_{i=1}^{n} l(i,k), \tag{2}$$

where *l(i, k)* is the length of the shortest path connecting *i*th and *k*th nodes.

Betweenness centrality [10] measures the extent to which a particular node lies between other nodes in a network. The betweenness of *k*th node is defined as the number of geodesics passing through it:

$$C_B(k) = \sum_{k}^{n} \sum_{j}^{n} g_{ij}(k), \tag{3}$$

where g_{ij} (*k*) indicates whether the shortest path between two other *i*th and *j*th nodes passes through the *k*th node. A member with high betweenness may act as a gatekeeper or broker in a network for smooth communication or flow of goods (e.g., drugs).

A more sophisticated version of the same idea is the so-called Eigenvector centrality. *Eigenvector centrality* x_i of a node in a network is defined to be proportional to the sum of the centralities of the node's neighbours, so that a node can acquire high centrality either by being connected to a lot of others (as with simple degree centrality) or by being connected to others that themselves are highly central. We write

$$x_i = \frac{1}{\lambda} \sum_{j=1}^{n} A_{i,j} x_j \, , \tag{4}$$

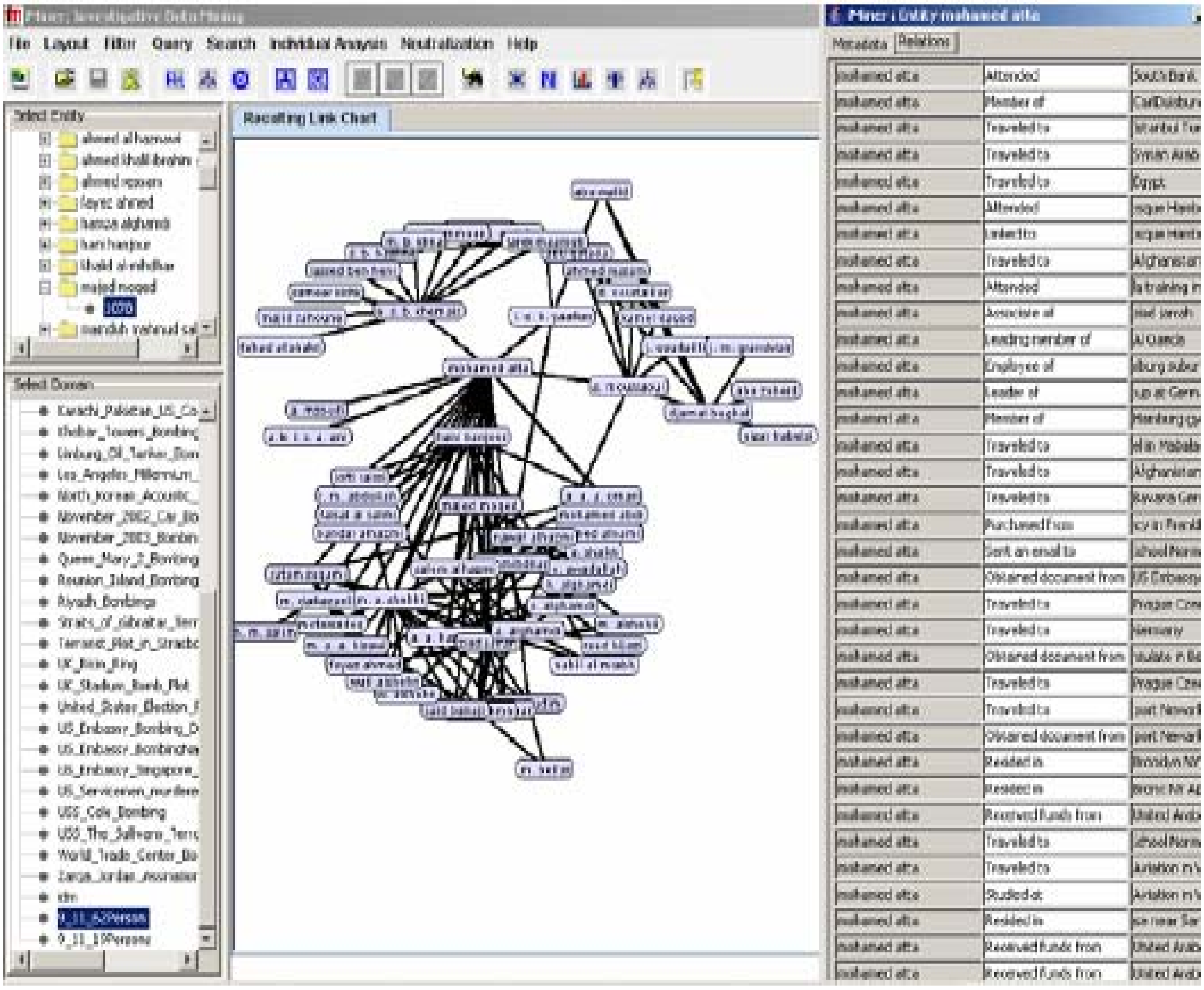
where n is total number of nodes and λ is constant. In matrix notation this becomes λ**x** = A**x**, so that x is an eigenvector of the adjacency matrix [11]. Assuming that we wish the centralities to be non-negative, it can be shown that λ must be the largest Eigen value of the adjacency matrix and **x** the corresponding Eigenvector.

2.1.1 Example

Figure 1 shows an example of a terrorist network, which maps the links between terrorists involved in the tragic events of September 11, 2001. This graph was constructed by Valdis Krebs [12] using the public data that were available before, but collected after the event. Even though the information mapped in this network is by no means complete, its analysis may still provide valuable insights into the structure of a terrorist organization. This graph is reconstructed in this paper, using metadata (additional information) of every terrorist involved in the attacks.

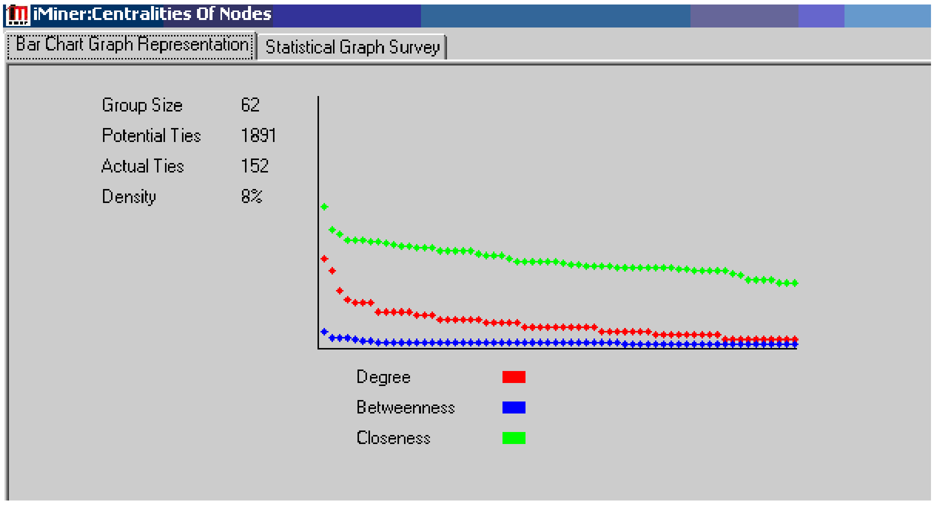
According to Kreb's analysis [12], this network had 62 members in total, of which 19 were kidnapers, and 43 assistants: organizers, couriers, financiers, scouts, representatives, coordinators, counterfeiters, etc. Allen [13] found that successfully functioning large networks typically comprise 25-80 members, with optimal size between 45 and 50. A close match exists between the results of Allen's analysis of collaborating networked groups and this particular example of terrorist group.

Inspection of this network by standard measures of network structure reveals firstly its low connectedness. A member of this network holds only 4.9 connections with others members on average (also known as degree centrality), which means that average members were rather isolated from the rest of the network. The density (which is defined as the number of actual links divided by the number of possible links) of this network is only 0.08, meaning that only 8% of all possible connections in the network really exist (see Figure 2).



mohamed atta	Attended	South Bank
mohamed atta	Member of	CalDuisburg
mohamed atta	Traveled to	Istanbul Tur
mohamed atta	Traveled to	Syrian Arab
mohamed atta	Traveled to	Egypt
mohamed atta	Attended	ogue Hamb
mohamed atta	Linked to	ogue Hamb
mohamed atta	Traveled to	Afghanistan
mohamed atta	Attended	la training in
mohamed atta	Associate of	Ziad Jarrah
mohamed atta	Leading member of	Al Qaeda
mohamed atta	Employee of	burg subur
mohamed atta	Leader of	up at Germ
mohamed atta	Member of	Hamburg gr
mohamed atta	Traveled to	Hilin Hotel
mohamed atta	Traveled to	Afghanistan
mohamed atta	Traveled to	Bavaria Ger
mohamed atta	Purchased from	cy in Florid
mohamed atta	Sent an email to	School Norm
mohamed atta	Obtained document from	US Embassy
mohamed atta	Traveled to	Prague Cze
mohamed atta	Traveled to	Germany
mohamed atta	Obtained document from	sulate in Be
mohamed atta	Traveled to	Prague Cze
mohamed atta	Traveled to	port Newark
mohamed atta	Obtained document from	port Newark
mohamed atta	Resided in	Brooklyn NY
mohamed atta	Resided in	Bronx NY A
mohamed atta	Received funds from	United Arab
mohamed atta	Traveled to	School Norm
mohamed atta	Traveled to	Aviation in V
mohamed atta	Studied at	Aviation in V
mohamed atta	Resided in	ice near Sar
mohamed atta	Received funds from	United Arab
mohamed atta	Received funds from	United Arab

Figure 1. The dataset of 9-11 hijackers and their affiliates. The dataset originally constructed by Valdis Krebs, but re-constructed in iMiner, using additional information of every entity.



Group Size	62
Potential Ties	1891
Actual Ties	152
Density	8%

Degree 🟥
Betweenness 🟦
Closeness 🟩

Figure 2. 9-11 Terrorists Network’s Neighbourhood

In spite of low connectedness, however, the nodes of this network are relatively close. The average closeness of nodes is 0.35. Betweenness as stated above is another important measure in SNA and it indicates a node's importance for communication among other nodes. The average betweenness of this network is 0.032, indicating relatively high average redundancy. However the betweenness of 40 nodes is in fact less than 1% and only 6 nodes have betweenness higher than 10%. These 6 nodes are critical for information flow, especially one with betweenness of almost 60%, meaning that almost 60% of communication paths among other nodes pass through this central node. The node represents *Mohamed Atta*, the leading organizer of the attack whose central position in the network is confirmed by other centrality indicators as well.

Distribution of degrees of nodes is particularly interesting. Degrees of nodes are exponentially distributed: the degree of most of the nodes is small, while few nodes have high degree (see Figure 2 and Figure 3). This property characterises the so called scale free networks [14] [15]. Scale free networks form spontaneously, without needing a particular plan or interventions of central authority. Nodes that are members of the network for a longer time, that are better connected with other nodes, and that are more significant for network's functioning, are also more visible to new members, so that the new members spontaneously connect more readily to such nodes than other, relatively marginal ones.

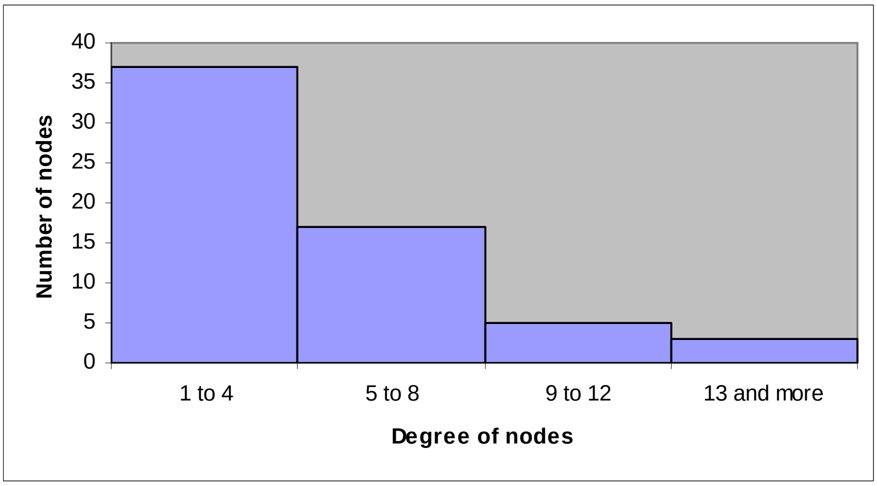


Figure 3. Distribution of degrees of nodes in the network (see Figure 1) of kidnappers and their supporters.

On the pattern of scale free networks, the Al Qaeda's Training Manual [16] states: "The cell or cluster methods should be organized in a way that a group is composed of many cells whose members do not know each other, so that if a cell member is caught, other cells would not be affected, and work would proceed normally".

2.2 Link Analysis

Link analysis is an analytic technique for making relationship explicit. The link analysis is the process of building up networks of interconnected objects through relationship in order to expose patterns and trends. Link analysis uses item-to-item associations to generate networks of interactions and connections from defined datasets. Link analysis methods add dimensions to an analysis that other forms of visualization do not support. Link analysis can be used to construct inferential structures of organizations or interactions which can be tested later. It is very well suited for hypothesis construction and can be applied to a variety of problems in Armed Forces Intelligence (for example, orders of battle), in Political Intelligence and in

Sociological Intelligence research and analysis [17].

Despite the seeming novelty of link analysis, the federal government in USA has used link analysis, for nearly fifty years. Karl Van Meter describes the two main types of the link analysis: the village survey method and traffic analysis [29]. The village survey method was created and used by Ralph McGehee of the CIA in Thailand in the 1960s to understand family and community relationships. He conducted a series of open-ended interviews and in a short time was able to map out clandestine structure of local and regional Communist organizations and associated sympathetic groups. Traffic analysis (also known as communication link analysis) began during World War II and its importance continues to this day. This technique consists of the study the external characteristics of communication in order to get information about the organization of the communication system. It is not concerned with the content of phone calls, but is interested in who calls whom and the network members, messengers, and brokers. Traffic analysis was also used by the British MI5 internal security service to combat the IRA in the 1980s and 1990s and continued to be used across the world by law-enforcement agencies including the US Defense Intelligence Agency (DIA) Office of National Drug Control Policy [29].

The *Analyst Notebook* [32] is the primary software used for link analysis. Currently on its sixth version, this software is recognized as one of the world's leading analytical tools and is employed in more than 1,500 organizations. SNA improves upon link analysis by moving from single variable analysis to multivariate analysis, allowing the individual to control many factors at once. The change from single variable to multivariate analysis is quite significant when researching terrorism: a number of factors affect terrorism, not one single factor. For example, the prediction for one to participate in a terrorist activity might not be strongly affected by the single variable of being related to a terrorist member. However, the combination of multiple variables, such as poverty and type of government, combined with the link to a terrorist member, may cause a person to participate in a terrorist activity. Multivariate analysis allows us to take into account these multiple variables and their effects when controlling another variable.

From the outside, it is difficult to understand how link analysis is being used in the federal governments. Confidentiality prevents government analysts from discussing their work with researchers and private companies without security clearances. Despite this lack of information, it is clear that government is interested in using network techniques in dealing with counterterrorism strategies. Many government agencies, such as Defense Advanced Research Projects Agency (DARPA), U.S. Army Research Labs, the U.S. Office of Naval Research (ONR), the National Security Agency (NSA), the National Science Foundation (NSF), and the Department of Homeland Security (DHS) have funded research related to link analysis.

The link analysis systems are being used in the investigative world, but from unclassified work, it is known that they are only used for identifying central players and some interesting patterns from the available datasets. Apparently little work is carried out for destabilizing of terrorist networks [1] [2] [4] [18] [26]. The motivation behind this study is to connect the dots in order to assist law enforcement agencies to disconnect / destabilize the most of the network by capturing/eradicating some key players.

3.0 EXISTING APPROACHES

Existing terrorist network research is still at its incipient stage. Although previous research, including a few empirical ones, have motivated the call for new approaches to terrorist network analysis [18] [19] [20], studies have remained mostly small-scale and used manual analysis of a specific terrorist organization. For instance, Krebs [12] manually collected data from public news releases after the 9/11 attacks and studied the network surrounding the 19 hijackers. Sageman [21] analyzed the Global Salafi Jihad network consisting of 171 members using a manual approach and provided an anecdotal explanation of the formation and evolution of this network. None of these studies used advanced data mining technologies that have been applied widely in other domains such as finance, marketing, and business to discover previously unknown patterns from terrorist networks.

The papers [18] [22] provide examples of social network analysis in counterterrorism applications and indicate both usefulness and some limitations of social network analysis as a basis for quantitative methods for situation awareness and decision-making in law enforcement applications. Raab & Milward [22] discuss the organizational structure of certain drug trafficking, terrorism, and arms trafficking networks, showing how some of them have adapted to increased pressure from the States and international organizations by decentralizing into smaller units linked only by function, information, and immediate need. They also describe ways and structures of cooperation between different kinds of criminal networks, for example in financing terrorists by illegal diamond and drug trafficking.

Recently, computer scientists have become interested in network analysis. This has led to an increased emphasis on studying the statistical properties of large networks, such as the Internet, criminal networks, and, even, infrastructure networks [23]. This influx of people to the field has also led to several new approximate algorithms to compute important properties [1] [2] [24] [25]. Most of the above mentioned literature has used graph theory and SNA techniques for terrorist network analysis.

Jonathan D. Farley presented a new mathematical approach [26] to destabilize terrorist networks using order theory. This paper pointed that modeling terrorist networks as graphs does not give enough information to deal with the threat, “modeling terrorist cells as graphs ignores an important aspect of their structure, namely their hierarchy, and the fact that they are composed of leaders and followers ". Jonathan D. Farley proposed an alternative approach that better reflects an organization's hierarchy. In this case, the relationship of one individual to another in a cell becomes important. Leaders are represented by the topmost nodes in a diagram of the ordered set representing a cell and foot soldiers are nodes at the bottom. Disrupting the organization would be equivalent to disrupting the chain of command, which allows orders to pass from leaders to foot soldiers [26].

What is needed is a set of integrated methods, technologies, models, and tools to automatically mine data and discover valuable knowledge from terrorist networks based on large volumes of data of high complexity.

The system to be proposed, we represent a terrorist network by an undirected graph; then we convert it into a directed graph with the help of centrality measures [1] [2]. Then we propose an approach for destabilizing the terrorist network to the converted directed graph into a hierarchical chart using dependence centrality [1]. From the hierarchical chart, investigators and law enforcement agencies can easily distinguish the leaders and peripheries in the network in order to destabilize the network. Our newly introduced dependence centrality measure may also be very useful in destabilizing terrorist networks, because it shows the nodes which are totally depending on particular nodes. If the nodes are completely depending on the other nodes, they will be isolated (cut-off from the network completely) by capturing the node on which those nodes are depending.

4.0 DESTABILIZING TERRORIST NETWORKS

In this section we present a theory behind analyzing and destabilizing of terrorist networks. We have implemented all the models discussed in this section in iMiner.

4.1 The Efficiency of a network

The network efficiency E (G) is a measure to quantify how efficiently the nodes of the network exchange information [27]. To define efficiency of a network G, first we calculate the shortest path lengths d_{ij} between i^{th} and j^{th} nodes. Let us now suppose that every node sends information along the network, through its links. The efficiency in the communication between i^{th} node and j^{th} node is inversely proportional to the shortest distance: when there is no path in the graph between i^{th} and j^{th} nodes, we get $d_{ij} = +\infty$ and efficiency becomes zero. Let N is known as the size of the network or the numbers of nodes in the graph, the average efficiency of the graph (network) of G can be defined as:

$$E(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in G} \frac{1}{d_{ij}} \tag{5}$$

The above formula gives a value of E in the interval of [0, 1].

4.2 Position Role Index (PRI)

The PRI is our recently introduced measure [2] which highlights a clear distinction between followers and gatekeepers (It is a fact that leaders may act as gatekeepers). It depends on the basic definition of efficiency as discussed in equation (5). It is also a fact that the efficiency of a network in presence of followers is low in comparison to their absence in the network. This is because they are usually less connected nodes and their presence increases the number of low connected nodes in a network, thus decreasing its efficiency.

$$\text{PRI} = \text{E (G)} - \text{E (G} - \text{v}_i) \text{ , i} = 1,...,\text{N}, \tag{6}$$

where G − v_i indicates the network obtained by deactivating node v_i in the graph G. If we plot the values on the graph, the nodes which are plotted below x-axis are followers, whereas the nodes higher than remaining nodes with higher values on positive y axis are the gatekeepers.

4.4 Dependence Centrality (DC)

The dependence centrality of a node is defined as how much that node is dependent on any other node in the network. It is defined as:

$$C_{\delta}(i) = \sum_{i \neq k, k \in G}^{n} \frac{m(i, j)}{N_p} + \Omega, \tag{7}$$

where *i*th node is the root node which depends on *j*th node, $\mathbf{N_p}$ is the number of geodesic (shortest path) from *i*th node to node *k*th through *j*th node. Also *m(i,j)* is inverse of geodesic distance *1/d(i,j)*)from *i*th node to *j*th node**.** The value of Ω is taken 1 if graph is connected and 0 in case it is disconnected. In this paper we take Ω as 1, because we consider that graph is connected.

The first part of the formula tells us that:

How many times *i*th node uses *j*th node to communicate with other *k*th node of the network? In simple words *k* is the node of the network, to which node *i* is connected through node ***j*.** (The connection represents the shortest path of node *i* to node *k*, and node *j* is in between).

We applied the above mentioned measures (described in subsection 4.1-4.2) in the network of terrorists involved in tragic events of September 11, 2001 (as shown in Figure 1). The results are depicted in Figure 4. The results show that node 33 (Mohamed Atta) as key player in the plot. The Δ E and PRI of this node is higher than all nodes which prove that this node played an important role in the plot and worked as gatekeeper and removing this node the efficiency of the graph is decreased from 0.395 to 0.32. This clearly identifies the importance of this node in the network. All the models presented in previous sections are implemented in iMiner.

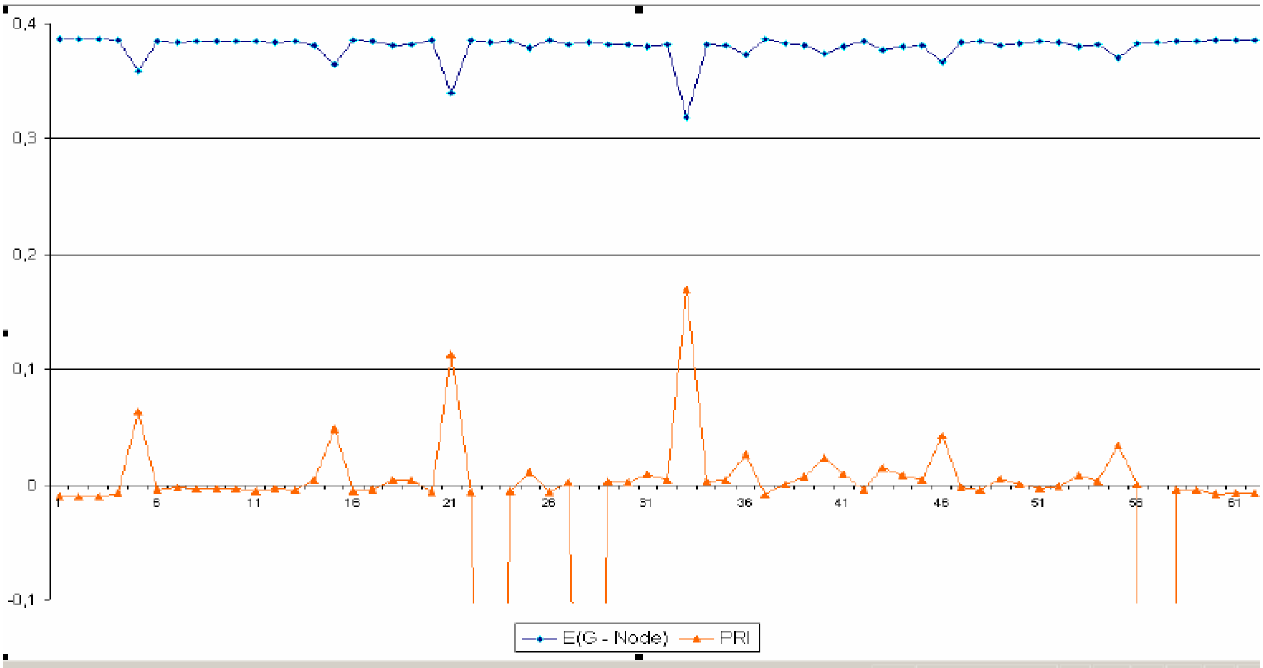


Figure 4. The efficiency of the original network E(G) = 0.395. The removed node is shown on x-axis. The efficiency of the graph once the node is removed is shown as E (G – v_i); while importance of node. The newly introduced measure position role index is shown as PRI.

5.0 THE SOFTWARE PROTOTYPE

The iMiner is an experimental system, which provides the answers of the following questions.

- Who is likely to be an important person (node) in the network?
- Why s/he is important?
- Which terrorist is highly/ less connected?
- What are various position roles in the network?
- Is there any command structure in the network?
- Is it possible to construct hidden hierarchy of the non-hierarchical networks?
- Which nodes represent key players?
- What is the efficiency of the network?
- How much the efficiency of the network is affected by eradicating one or more terrorists?
- How can the law enforcement agencies use (often incomplete and faulty) network data to disrupt and destabilize terrorist networks?

The system architecture of investigative data mining toolkit (iMiner prototype) is depicted in Figure 5. The first stage of network analysis development is intended to automatically identify the strongest association paths, or geodesics, between two or more network members. In practice, such task often entails intelligence officials to manually explore links and try to find association paths that might be useful for generating investigative leads.

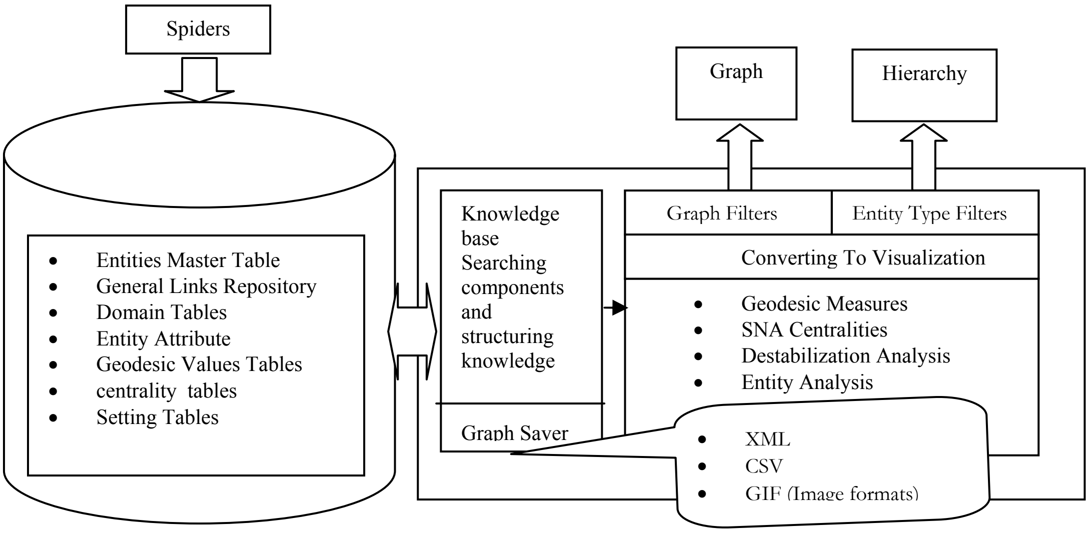


Figure 5. The System Architecture of iMiner

The iMiner allows the analyst to determine hierarchy of covert networks, which may help law enforcement and intelligence agencies in understanding the structure of the covert networks. In addition to this, intelligence agencies can easily set the priorities for eradicating some important nodes in the network by visualizing how much the efficiency of the network is minimized by the capture of a node.

The iMiner knowledgebase has following type of tables

- Entities Master Table
- General Links Repository
- Domain Tables
- Entity Attribute
- Geodesic Values Tables
- centrality tables
- Setting Tables

The above tables are briefly described below:

Entities Master Table
This table is a table where entity is assigned with unique id and very primary information about entity like entity's name and type being saved.

General Links Repository
This table has a record of all relationships in form of Association Data Model. These all relationships are used while searching the database and graphing a particular plot.

Domain Tables
These tables are subsets of general links repository. The relationships which are connected to a particular domain are stored separately in that domain table so that the complex operations which performed by iMiner during analysis takes place efficiently.

Entity Attribute Tables
These tables are used to store attributes of entities.

Geodesic Values Tables
These tables are used to save the geodesic values which are calculated during analysis of a particular domain and are useful for other analysis activities. The iMiner has components building to show those values in anytime to help analysts.

Centrality Tables
These tables are used to save the values of centralities (SNA centralities, dependence centrality, position role index, etc.) of each node in a graph of a particular domain under investigation.

Setting Tables
There are some tables which are used to save settings of iMiner, like search depth, entity types along with their visualization scheme, etc.

In the aftermath of the September 11th attacks, it was noted that coherent information sources were not available to the researchers [3]. Information was either available in fragmentary form, not allowing comparison studies across incidents, groups or tactics, or made available in written articles – which are not readily suitable for quantitative analysis of terrorist networks. Data collected by law enforcement agencies, while potentially better organized, are largely not available to the research community due to restrictions in distribution of sensitive information.

To counter the information scarcity, we have developed knowledgebase about the terrorist attacks occurred in the past and the information about terrorist organizations involved in those events. This information is mostly collected from open source media (but authenticated websites), such as http://www.trackingthethreat.com/.

The focus of the knowledgebase we have developed is the agglomeration of publicly available data and integration of the knowledgebase with investigative data mining software prototype. The main objective is to investigate and analyze terrorist networks to find hidden relations and groups, prune datasets to locate regions of interest, find key players, characterize the structure, trace point of vulnerability, and detect efficiency of the network and to discover the hidden hierarchy of the non-hierarchical networks. The Website http://www.trackingthethreat.com/ represents the original open-source database on Al Qaeda. It contains data in the form of:

Entities: Discrete data elements that comprise people, places, organizations, events, etc.

Relationships: Information about the personal, organizational, transactional, and historical connections between entities.

Metadata: Additional information about entities and relationships that help form a more complete picture.

Notes and Documents: Unstructured text that provides background information on entities, relationships, and metadata.

The iMiner system applies Spider Management System as well as spiders to import data that is available at World Wide Web. It can be developed to get information from online repositories and save it in its knowledgebase for analysis as shown in Figure 6.

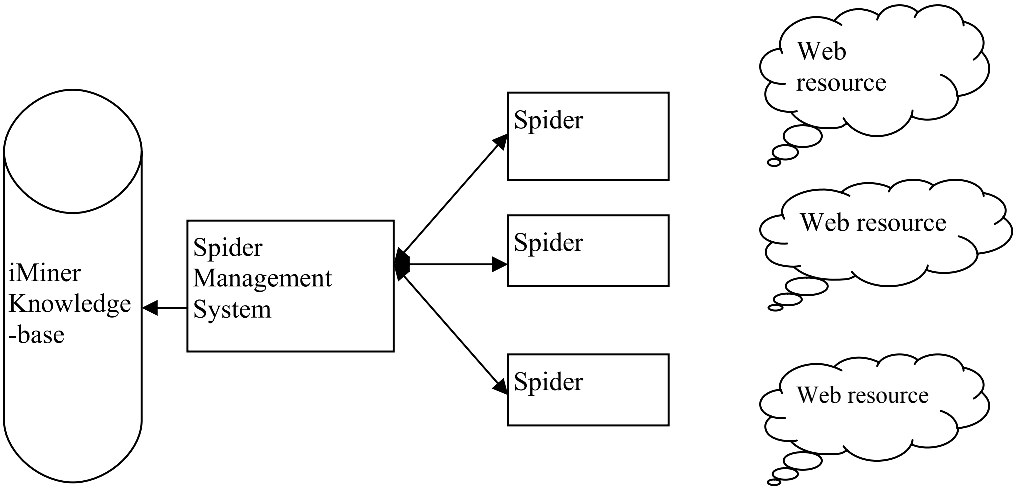


Figure 6. Spidering the information from Internet

The iMiner spidering system has spider implementation for a web resource as mentioned above (http://www.trackingthethreat.com/). Similarly this model can be evolved to make most of information present in shape of static and dynamic websites. The data flow model for the prototype is shown in Figure 8.

The iMiner supports domain-based study of network system. The domain can be similar to a particular case study (for example Bali Bombing or 9/11 case studies) with resizable boundaries. An investigating officer may expand the domain by including additional entities or reduce it by excluding less important or partially involved entities. The main advantage of this type of approach is to isolate a piece of data in which an investigating officer is interested, to concentrate analysis activities only on those entities.

To reuse the type of data, the iMiner has a data transform subsystem; the purpose of the subsystem is to import data from different legacy systems and transform them into the format required by iMiner knowledge base to carry out further investigation as shown in Figure 7.

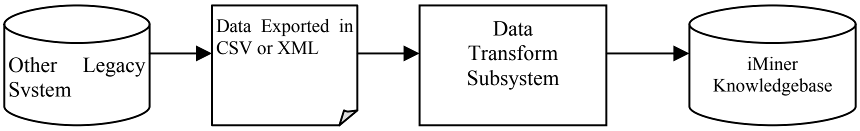


Figure 7. The data transform sub-system

Most of the commercial database solutions provide facilities to export data in at least CSV or XML format. The information exported from any legacy system or data source in CSV or XML format can be imported to extend iMiner knowledgebase to proceed with further investigation.

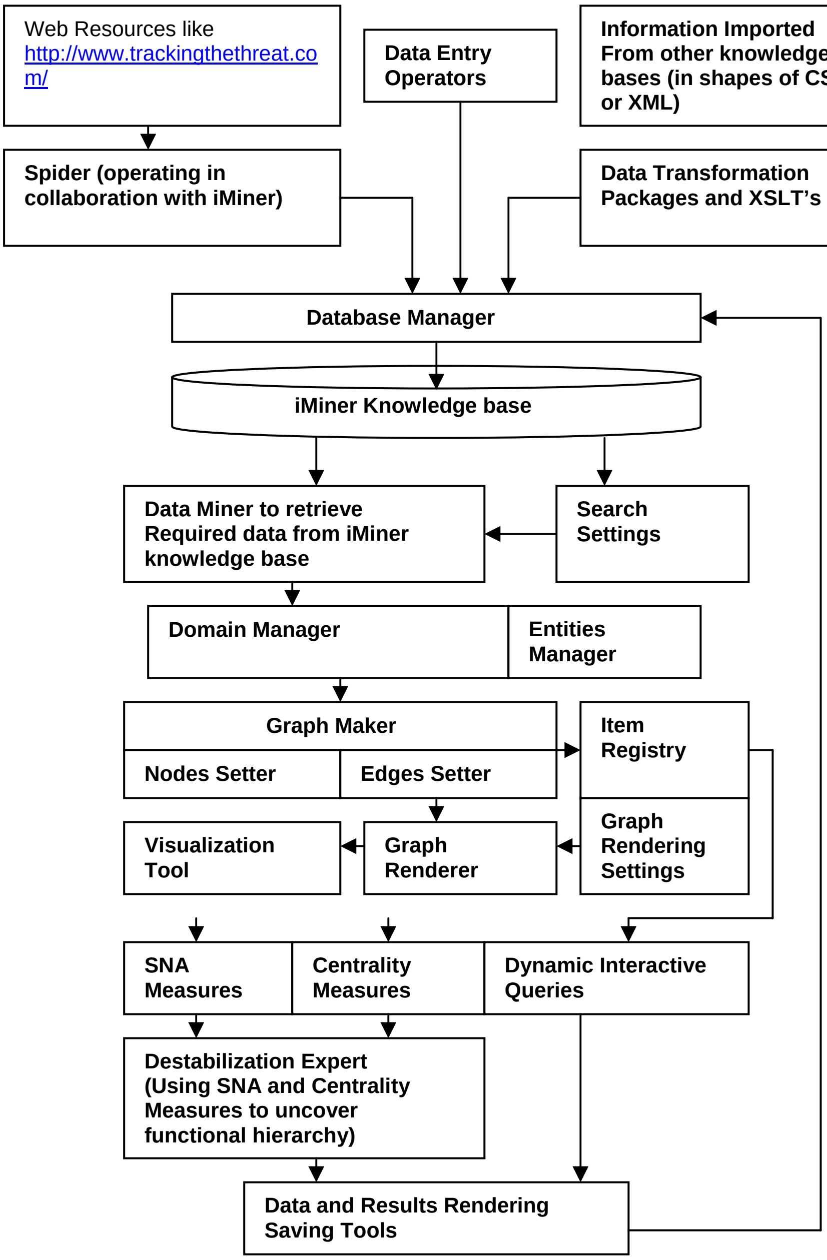

Web Resources like http://www.trackingthethreat.com/
Data Entry Operators
Information Imported From other knowledge bases (in shapes of CSV or XML)
Spider (operating in collaboration with iMiner)
Data Transformation Packages and XSLT's
Database Manager
iMiner Knowledge base
Data Miner to retrieve Required data from iMiner knowledge base
Search Settings
Domain Manager
Entities Manager
Graph Maker
Nodes Setter
Edges Setter
Item Registry
Visualization Tool
Graph Renderer
Graph Rendering Settings
SNA Measures
Centrality Measures
Dynamic Interactive Queries
Destabilization Expert (Using SNA and Centrality Measures to uncover functional hierarchy)
Data and Results Rendering Saving Tools

Figure 8. The data flow model of iMiner

6.0 CASE STUDIES

In this section, we present four case studies of the terrorist attacks occurred / planned in the past to test the algorithms and models recently published [1] [2].

6.1 Bali Night Club Bombing Terrorist Attack

On October 12, 2002 a car bomb exploded outside the Sari night club in Bali, a popular tourist island in Indonesia [28]. The attack was the worst terrorist incident in the history of Indonesia, with 202 civilians dead and more than 100 wounded. According to the Australian Broadcasting Corporation, the mastermind, behind the attacks was Al-Qaeda's chief representative and senior planner in Southeast Asia, as well as being operational chief of Jemaah Islamiya, Ryuduan bin Isomuddin, also known as Hambali. Hambali was detained by the U.S. government in August of 2003. Hambali was also believed to have been involved in the 2003 Marriot Hotel bombings in Jakarta, facilitated the January 2000 meeting in Malaysia including two September 11th hijackers, and an associate of 9/11 mastermind Khalid Sheikh Mohammed.

The Bali Night Club Bombing 2002, when searched in iMiner knowledgebase, results the graph as shown in Figure 9. Using algorithms for destabilizing terrorist networks (for details see [2]), we succeeded in constructing the hidden hierarchy of Bali bombing terrorist attack, shown in Figure 10, using iMiner. This hierarchy has some unconnected nodes, where as you can find a hint of patterns some time. The H. B. A. Haq and its descendants form a group (This cluster was acted as executive cluster), while the cluster Khalid Sheikh Mohammed (and his affiliates) was well known as a strategic cluster, whereas Ryuduan bin Isomuddin (known as Hambali) and his associates cluster known as tactical / logistic cluster . The accuracy of the software can be determined by the fact that all of H. B. A. Haq, Khalid Sheikh Mohammed and R. Isomuddin were key players in the reality. H. B. A. Haq was termed as potential leader while Khalid Sheikh Mohammed was the key conspirator.

In this case study we consider the connections network of terrorist involved in Bali Night Club Bombing and their direct or indirect relationships with other entities. Of course, mapping networks after an event is relatively easy, while the real problem in this case is to map the covert networks to prevent terrorist activity, a task that can be more difficult. The network reported in Figure 9 is constructed by iMiner, using publicly released information taken from major newspapers/ websites. The network size (say N) is 125 nodes, actual ties (arcs/ links) are 195, potential ties, which can be calculated by using the formula:
N (N-1) / 2, are 7750. The density of the network (the number of actual ties divided by the number of potential ties) is approximately 2.5%. To individuate the critical nodes, *i.e.*, the terrorists who played key role in the network, we deactivate one by one (from the network, remember the size of the network is 125), then we calculate the efficiency of the new network and the drop of efficiency caused, discussed in section 4. In Table 1 we show 6 most important nodes ranked according to the measure position role index defined in section 4.2, the degree of each node (represented as k) is also reported.

Table 1. Effect of the deactivation of a node in the terrorist network of Bali Night Club Bombing 2002 attacks

Removed Node	$E(G-v_i)$	k	PRI
K. S. Mohammed	0.3063	27	0.1854
Riduan Isamudin	0.3329	23	0.1148
Yazid Sufaat	0.3342	12	0.1112
Wan Min	0.3700	11	0.0161
Huda Bin A. Haq	0.3745	12	0.0042
Osama Bin Laden	0.3745	3	0.0027

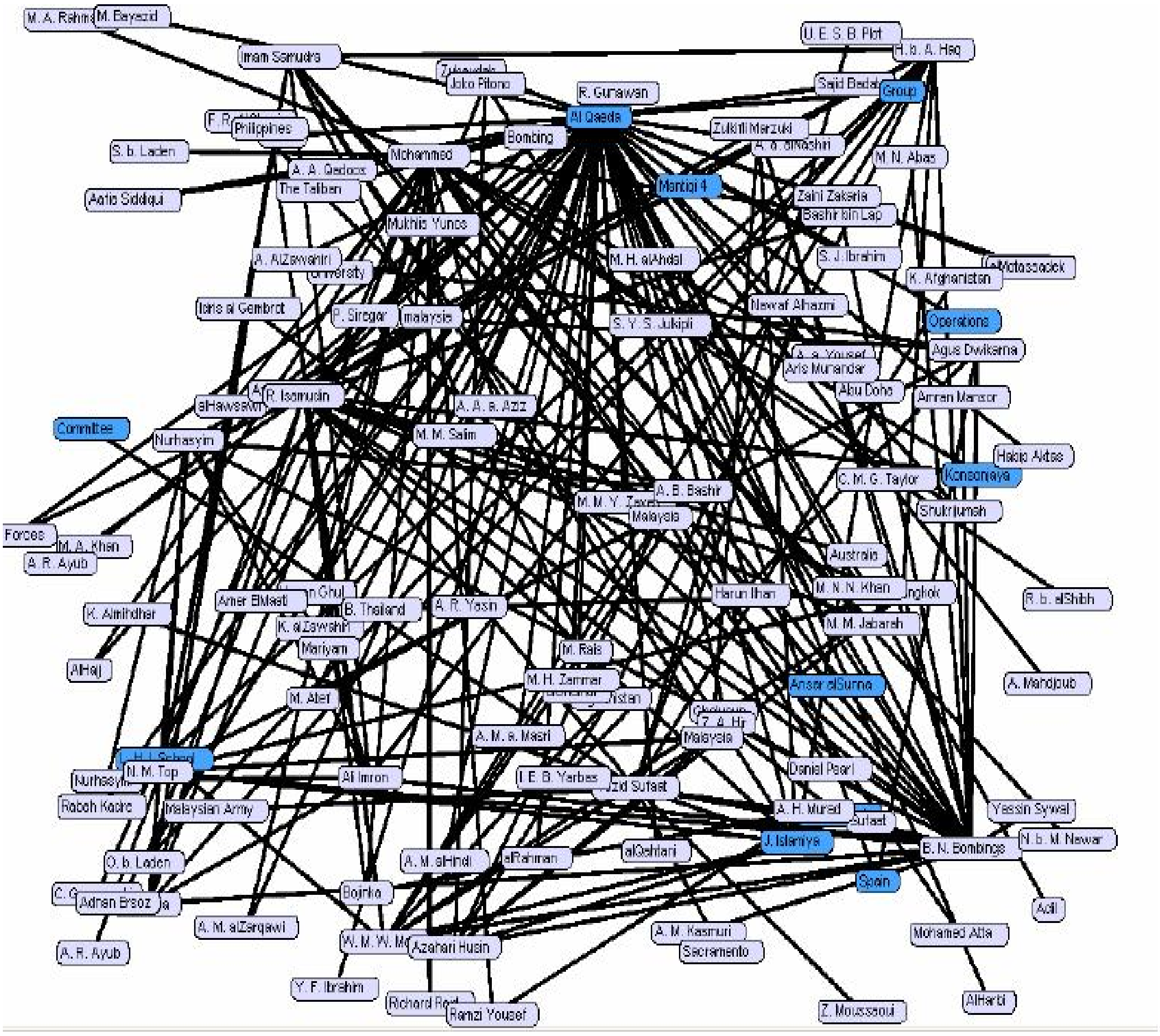


Figure 9. Terrorists involved in Bali Night Club Bombing attack and their directed and undirected relationships

The efficiency of the network is 0.382. The removed node is listed in the first column; the efficiency of the graph once the node is removed is reported in the second column, while the relative drop of efficiency is in the third column. In the fourth column we report, as an alternative measure of the importance of a node, the degree of (i.e. the number of links incident with) the removed node. The last column is reported with PRI of the nodes. The six most important nodes of the network are ordered according to their positions in the network, i.e. leaders/ gatekeepers/ followers.

It is crystal clear from the Table 1, that *K.S. Mohammed* was the most connected node in the network (k=27), and our recently introduced measure i.e. PRI also proves him as an important entity (*key player*), removing this node the efficiency of the network is decreased to 0.306. Moreover the measure PRI of *Huda Bin A. Haq* is lower than other leaders of two clusters, shows that he was the leader execution cluster. It is also interesting to note that in this network *Osama Bin Laden* is less connected node (k=3) and position role index measure shows his leading capabilities. Moreover, 64% of the entities in this network (80 out of 125) having degree (k) is equal to 1 (it means about 2/3rd of the network members degree is 1), entity Al Qaeda is the most connected node having k = 52 (it means that 52 nodes have directly affiliated with the terrorist organization Al Qaeda) and 29 nodes are directly connected with Bali bombing terrorist plot.

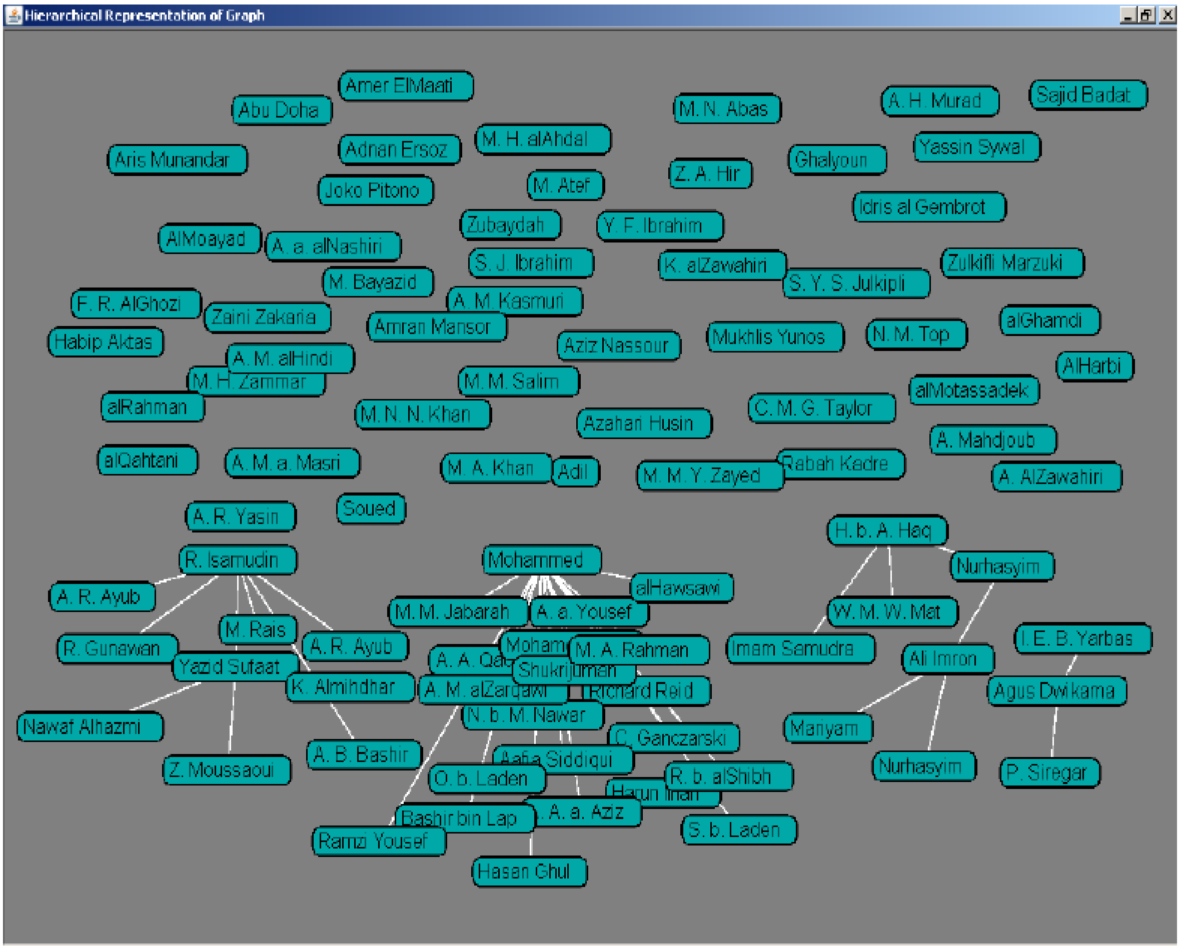


Figure 10. Hidden hierarchy constructed by iMiner for Bali Bombing Attack

6.2 Dirty Bomb Plot

Adnan Gulshair el Shukrijumah is a suspected of plotting to carry out a terrorist attack against U.S. interests abroad or domestically [30]. When Khalid Sheikh Mohammed, former operational commander of Al Qaeda, was captured and interrogated, he fingered Adnan el Shukrijumah as the man who would later be in charge of new attack. Shukrijumah is believed to be working on Osama bin Laden's plan to trigger a radiological disaster inside the United States [33] – the so-called "dirty-bomb" scenario where a small charge would trigger dispersion of radiation over a large area, wreaking havoc on those caught in the blast and making the blast area uninhabitable. High-grade uranium is not necessary for this project; ordinary, low-grade nuclear waste will be deadly enough.

El Shukrijumah has eluded capture. According to an FBI informant, El Shukrijumah was spotted 2002 in Hamilton, Ontario, posing as student at McMaster University, which has 5-megawatt research reactor. The U.S. officials believe El Shukrijumah was in Hamilton to obtain radioactive material. We have collected dataset for dirty bomb terrorist network plot which is depicted in Figure 11.

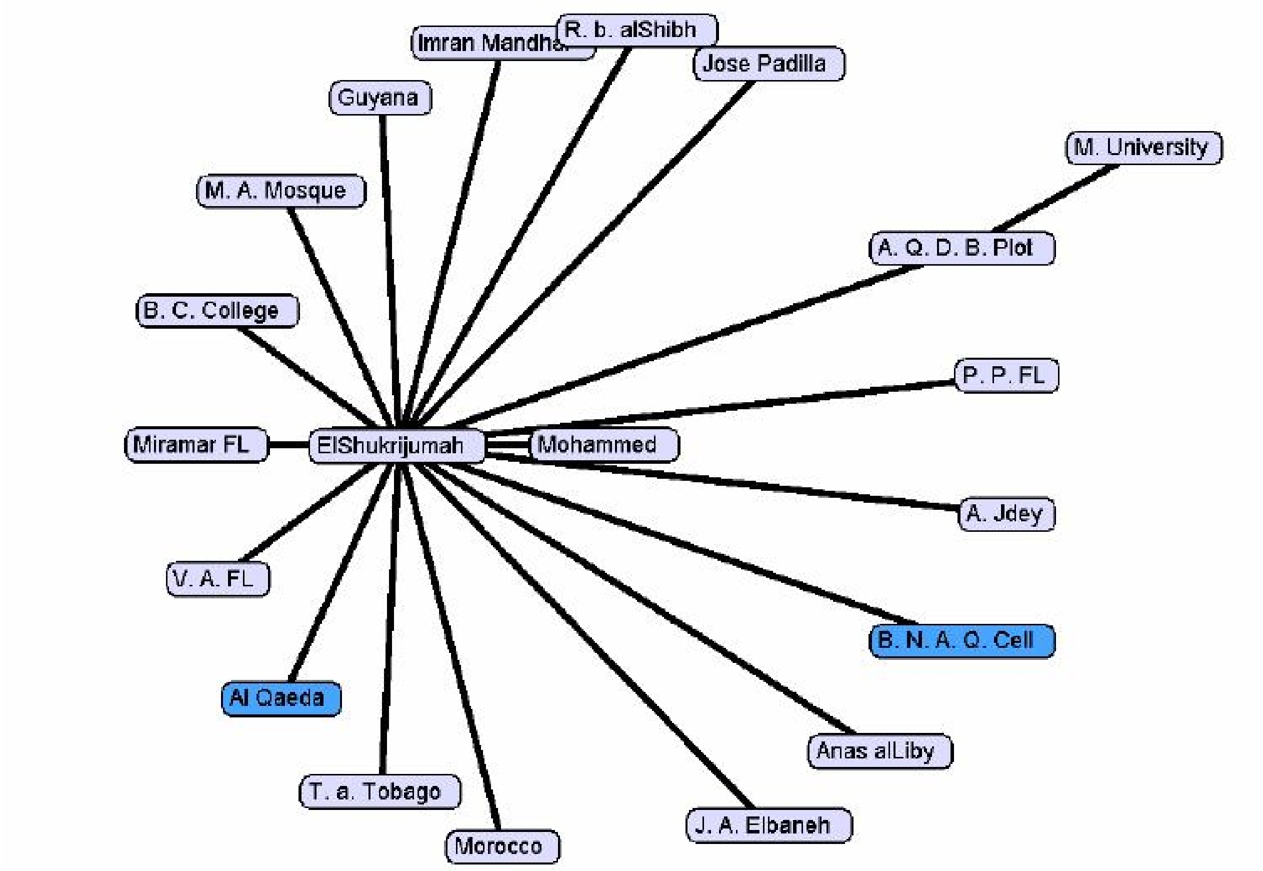


Figure 11. Terrorists' involvement for Dirty Bomb plot

Using the algorithms for construction of hidden hierarchy [2], the hierarchy for the dirty-bomb plot is shown in Figure 12.

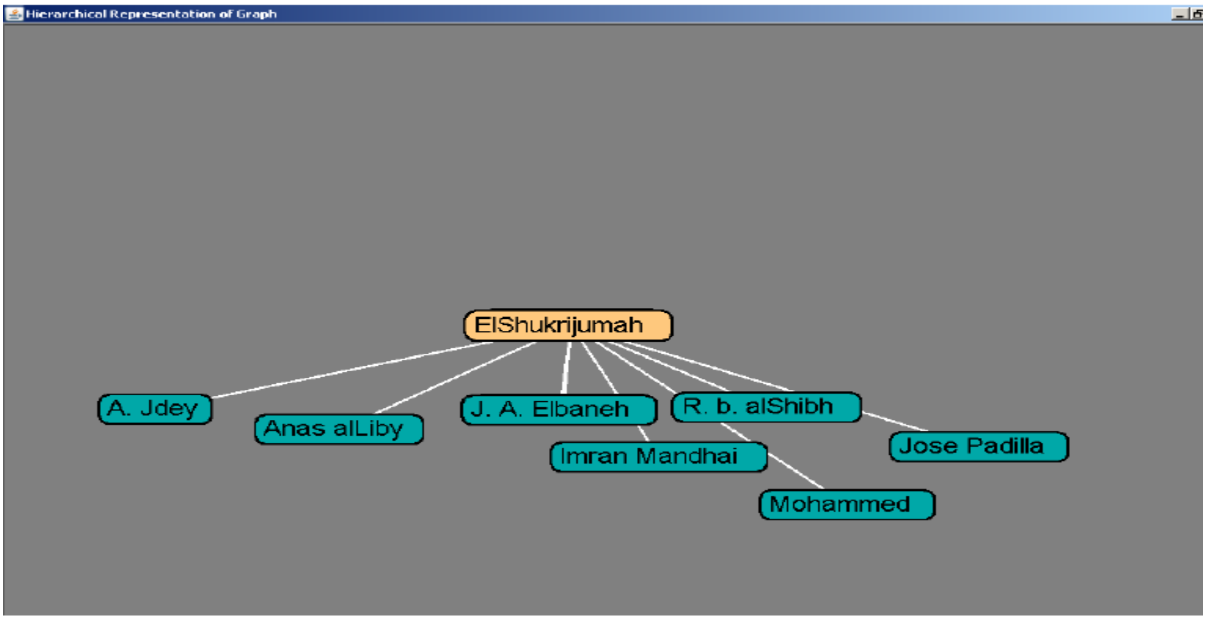


Figure 12. Hidden hierarchy constructed by iMiner for Dirty-Bomb Plot

It is believed by authorities that Shukrijumah may have been trained at an Al Qaeda terrorist camp[34]. Shukrijumah has extensive flight training that he received at a flight school in Florida and is a pilot, though he is not registered with the Federal Aviation Administration (FAA). Obviously, this is of concern to law enforcement since virtually all the Al Qaeda hijackers involved in the September 11 terrorist attacks, received training to be pilots at U.S. private flight schools.

Also, it is believed by the FBI that Shukrijumah has been trained by Al Qaeda to operate as a terrorist organizer and operational / field commander and lead or coordinate a terrorist assault, much the same way Mohammed Atta was designated and trained as an organizer and operational/field leader by Al Qaeda to lead the 9/11. El Shukrijumah may play a large and leading role in the next set of terrorist attacks to come upon the U.S. Shukrijumah was last seen in the Miami or southern Florida area in the early part of 2003. He has not been seen since and no one knows of his whereabouts [34].

6.3 WTC 1993 Bombing Plot

The WTC bombing attack in the garage was occurred on February 26, 1993 of the New York World Trade Center. A car bomb was planted by terrorist group in the underground parking garage below tower One. It killed six, injured over 1,000 and presaged the 9/11 attacks on the same buildings.

We have collected data of the terrorists involved in the attack and the network is shown in Figure 13.

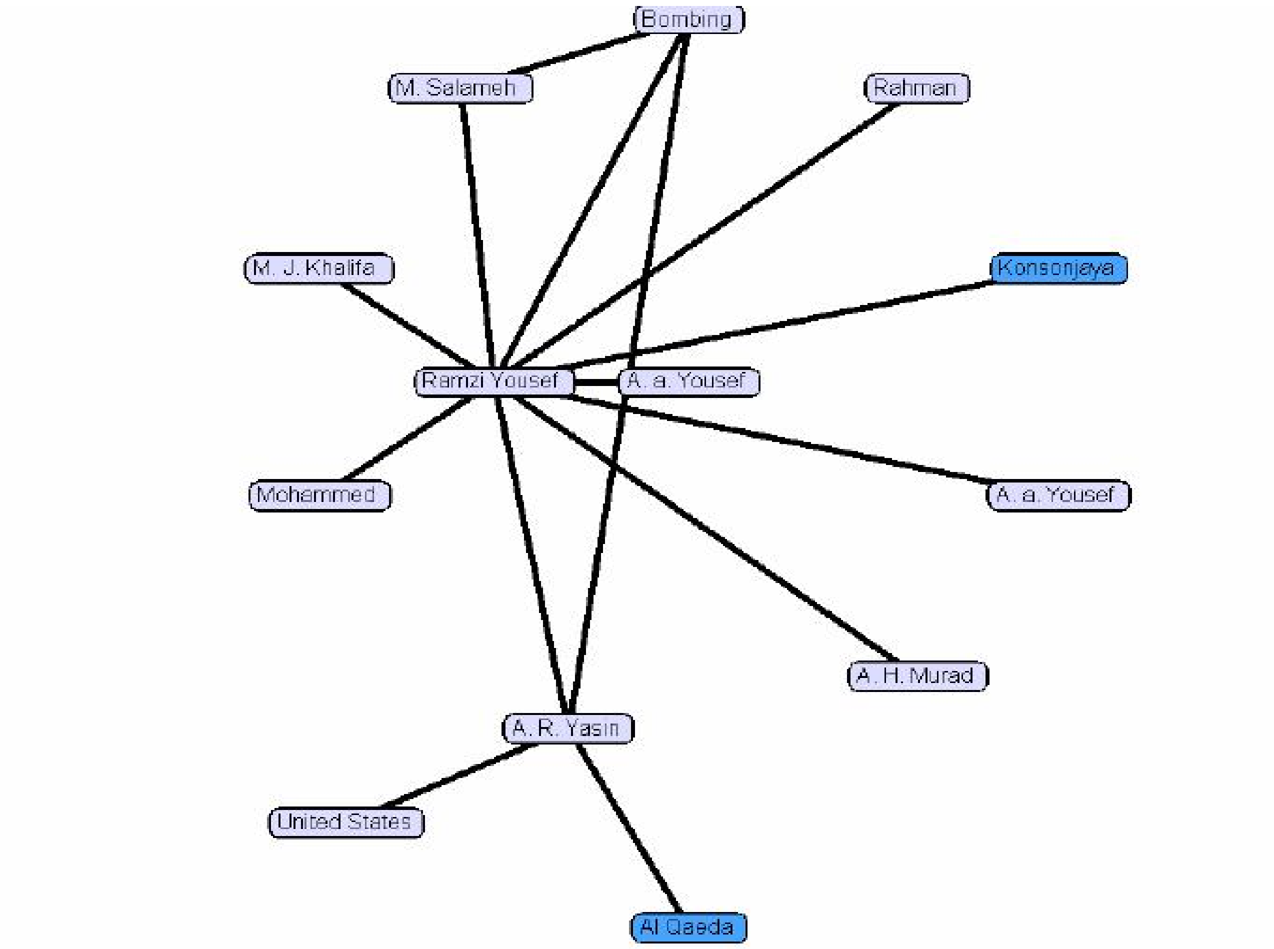


Figure 13. Terrorists Network involved in WTC 1993

The group size of the network is 13; potential ties are 78; whereas actual ties are 14. The density of the network is 18% and efficiency of the network is 0.555. The most important node in the network is Ramzi Yousef (k=10 and Eigenvector centrality = 16). Removing this node the efficiency of the graph is decreased to 0.0845.

Using algorithms for the construction of hidden hierarchy of covert/ terrorist networks [2], the hierarchy of the network is shown in Figure 14.

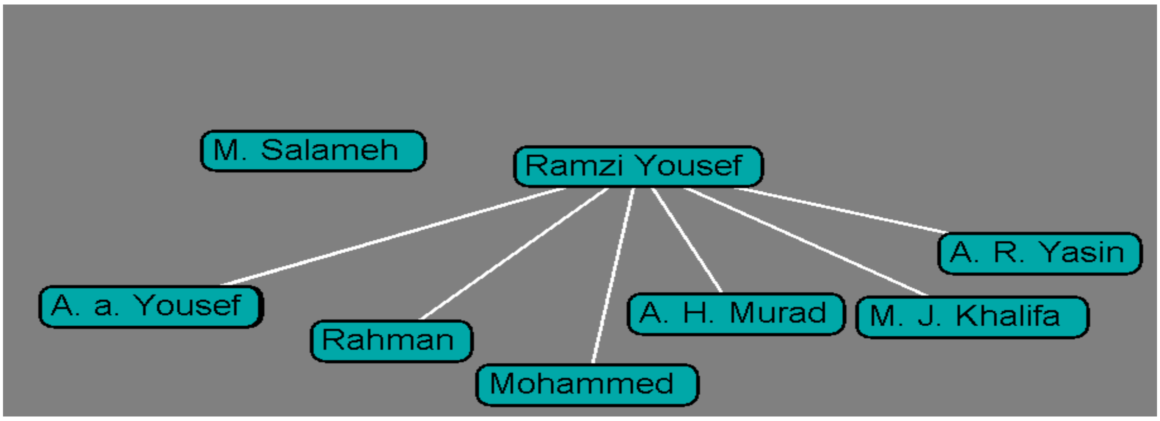


Figure 14. Hidden hierarchy constructed by iMiner for the terrorist network shown in Figure 13

It is fact that Ramzi Yousef began in 1991 to plan bombing attack with USA. Yousef's Uncle Khalid Sheikh Mohammed gave him advice and tips over the phone, and funded him. Yousef entered USA with a Iraqi passport without a U.S. visa on September 1, 1992 [35]. Muhammad Jamal Khalifa, Omer Abdel Rehman, Abdul Rehman Yasin, Abdul Hakim Murad were key companions of Ramzi Yousef. There is an error found in the dataset that Ahmed Yousef is an alias used by Ramzi Yousef, but mistakenly it appears as an independent actor (This is an error). Moreover M. Salameh was an informant and he informed FBI about the plot.

6.4 September 11, 2001 Terrorist Plot

The September 11, 2001 attacks (often referred to as 9/11—pronounced "nine eleven") consisted of a series of coordinated terrorist suicide attacks upon the United States, predominantly targeting civilians, carried out on Tuesday, September 11, 2001 [36].

That morning, 19 terrorists affiliated with al-Qaeda hijacked four commercial passenger jet airliners. Each team of hijackers included a trained pilot. Two aircraft (United Airlines Flight 175 and American Airlines Flight 11) crashed into the World Trade Center in New York City, one plane into each tower (WTC 1 and WTC 2). Both towers collapsed within two hours, followed by WTC 7 later that day. The pilot of the third team crashed American Airlines Flight 77 into the Pentagon in Arlington County, Virginia. Passengers and members of the flight crew on the fourth aircraft (United Airlines Flight 93) attempted to retake control of their plane from the hijackers; that plane crashed into a field near the town of Shanksville in rural Somerset County, Pennsylvania. As well as the 19 hijackers, a confirmed 2,973 people died and another 24 are missing but presumed dead as a result of these attacks [36].

The renowned Social Network Analyst Valdis Krebs [12] mapped the network of 9/11 (hijackers and their affiliates) as shown in Figure 1. Using the algorithms for detecting hidden hierarchy of non-hierarchical terrorist networks [2], we tested the network of terrorists involved in 9-11 tragic events and results are depicted in Figure 15.

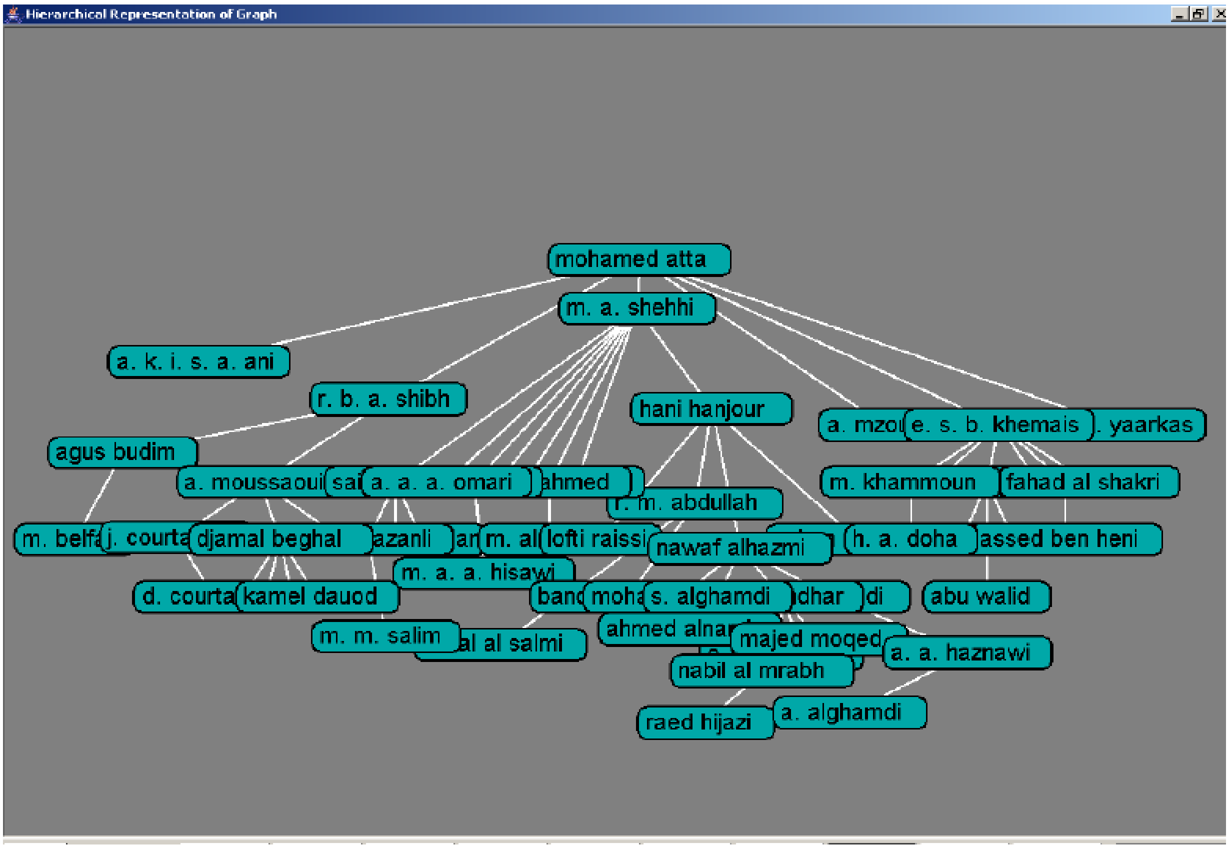


Figure 15. The hierarchy clearly suggests that Muhammad Atta was the most powerful person (leader) of the plot. While M.A. Shehhi was assisting him, as he is below in the hierarchy. They both were found the key leaders of the plot by 9-11 commission.

7.0 CONCLUSION

In this paper, we presented an overview of an investigative data mining toolkit (iMiner software prototype) which we have developed for undertaking analysis of terrorist networks. In general investigative data mining has been shown to be a promising and potentially powerful area of research. The paper presented interesting patterns gleaned from the data. We discussed three innovative ideas of our research which were already published in [2] and featured in Government Computer News [31]. The mathematical models and algorithms discussed in the paper are implemented in the prototype. The *iMiner* demonstrates key capabilities and concepts of a terrorist network analysis toolkit. Using the toolkit investigating officials can predict overall functionality of the network along with key players. Thus counterterrorism strategy can be designed keeping in the mind that destabilization not only means disconnecting the dots (nodes) but disconnecting those key players from the peripheries by which maximum network could be disrupted. The investigative mining can be used to understand terrorist networks, and we are of the view that investigative data mining tool like iMiner, score over traditional analysis of networks with large volume of data and investigative data mining could reduce the consequent overload on analysts. The results presented in this paper are our findings based on limited exercise in exploring the utility of investigative data mining in analyzing terrorist networks. The may also be used for law enforcement agencies for destabilizing of terrorist networks for capturing the key nodes. The intelligence agencies may also evaluate the efficiency of the networks by capture of a particular node.

Further real-time or near real-time information from multiplicity of databases could have the potential to generate early warning signals of utility in detecting and deterring terrorist attacks. It is necessary, of course, to have ‘experts’ in the loop. This analysis has provided substantive and in-depth analysis of terrorist networks. Furthermore this analysis has provided a richer and deeper understanding and insight into terrorist networks and has provided approaches to destabilize the networks.

In this paper we presented the system architecture of our software prototype and the process by which we harvested data from web and stored in the knowledgebase. The focus of the knowledge base we have developed is the agglomeration of publicly available data and integration of the datasets with the software prototype in order to investigate interested patterns. The flow model of the prototype is also shown and discussed

8.0 REFERENCES

[1] Memon Nasrullah and Henrik Legind Larsen.: *Practical Approaches for Analysis, Visualization and Destabilizing Terrorist Networks*. In the proceedings of ARES 2006: The First International Conference on Availability, Reliability and Security, Vienna, Austria, IEEE Computer Society, pp. 906-913, (2006)

[2] Memon Nasrullah and Henrik Legind Larsen: *Practical Algorithms of Destabilizing Terrorist Networks*. In the proceedings of IEEE Intelligence Security Conference, San Diego, Lecture Notes in Computer Science, *Springer-Verlag*, Vol. **3976**: pp. 398-411 (2006)

[3] Memon Nasrullah and Henrik Legind Larsen: Detecting Terrorist Activity Patterns using Investigative Data Mining Tool. *International Journal of Knowledge and System Sciences*, Vol. **3**, No. 01, pp. 43-52 (2006)

[4] Memon Nasrullah, Abdul Rasool Qureshi..: Destabilizing Terrorist Networks. *In WSEAS Transactions on Computers*, Issue 11, Vol. **4,** pp.1649-1656 (2005)

[5] Taskar Ben, Pieter Abbeel, Ming-FaiWong, and Daphne Koller: *Label and Link Prediction in Relational Data*, in IJCAI Workshop on Learning Statistical Models from Relational Data, (2003) http://kdl.cs.umass.edu/srl2003_upload/files/taskar-paper.pdf

[6] M. Barlow, J. Galloway, and H. Abbass, *Mining Evolution through Visualization*. In Proceedings of Workshop on Beyond Fitness: Visualization Evolution at the 8th International Conference on the Simulation and Synthesis of Living Systems, (2002) http://www.alife.org/alife8/workshops/15.pdf

[7] Q&A with Professor Karen Stephenson, April 18, 2006 http://www.elearningpost.com/articles/archives/qa_with_professor_karen_stephenson/

[8] Klerks, P. The Network Paradigm Applied to Criminal Organisations: Theoretical nitpicking or a relevant doctrine for investigators? Recent developments in the Netherlands. *Connections* 24(3): 53-65, (2002)

[9] Fijnaut Cyrille J.C.F., Frank Bovenkerk, Gerben Bruinsma. Organized Crime in Netherlands. The *Hauge Kulwar Law International*, (1998)

[10] Freeman, L.C. Centrality in Social Networks: I. Conceptual clarification. *Social Networks*, **1**:215-39 (1978)

[11] Bonacich, P. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology*. **2**:113-120 (1972)

[12] Krebs, V. E. Mapping Network of Terrorist Cells. *Connections* 24(3): 43-52 (2202)

[13] Allen, C. *The Dunbar Number as a Limit to Group Sizes* (2004). Retrieved May 31, 2006, from http://www.lifewithalacrity.com/2004/03/the_dunbar_numb.html

[14] Watts, D.J.: *Six Degrees – The Science of a Connected Age*. W.W. Norton & Company, New York, 2003.

[15] Kreisler, H.: International Relations in Information Age: Conversation with John Arquilla, Professor at University of California, Berkeley 2003. Retrieved on June 01, 2006 from http://globetrotter.berkeley.edu/people3/Arquilla/arquilla-con0.html

[16] Al Qaeda Training Manual. Retrieved on June 01, 2006 from

http://www.fas.org/irp/world/para/manualpart1.html

[17] Walter R. Harper and Douglas H. Harris, The Application of Link Analysis to Police Intelligence," *Human Factors* 17: 157-164 (1975)

[18] Carley, Kathleen Ju-Sung Lee, David Krackhardt. Destabilizing Networks, *Connections* **24**(3):31-34. (2001)

[19] McAndrew, D.: *Structural Analysis of Criminal Networks. Social Psychology of Crime: Groups, Teams, and Networks, Offender Profiling Series*, III. L. Allison. Dartmouth, Aldershot (1999)

[20]Sparrow, M. K.: Application of Network Analysis to Criminal Intelligence: An Assessment of the Prospects. *Social Networks* 13 , 251-274 (1991)

[21] Sageman, M.: Understanding Terror Networks. Pennsylvania, University of Pennsylvania Press (2004)

[22] J. Raab and H. B. Milward. Dark Networks as Problems. *Public Administration Research and Theory*, 13(4): 413-439 (2003)

[23] P. Svenson, C. Mårtenson, C. Carling. *Complex Networks: Models and Dynamics*, FOI-R-1766-SE (2005)

[24] A. Clauset, M. Newman, C. Moore. Finding Community Structure in Very Large Networks, *Physical Review* E70, 066111 (2004)

[25] M. Newman. A Measure of Betweenness Centrality based on Random Walks. *Social Networks* 27: 39-54 (2005)

[26] Farley, J.D. Breaking Al Qaeda Cells: A Mathematical Analysis of Counterterrorism Operations (A Guide for Risk Assessment and Decision Making). *Studies in Conflict and Terrorism*, 26: 399-411 (2003)

[27] Latora, V., Massimo Marchiori, How Science of Complex Networks can help in developing Strategy against Terrorism, *Chaos, Solitons and Fractals* 20, 69-75 (2004)

[28] Bali Night Club Bombing. Retrieved on June 01, 2006 from
http://www.globalsecurity.org/security/ops/bali.htm

[29] Karl M. Van Meter, "Terrorists/Liberators: Researching and Dealing with adversary social networks," *Connections* **24** **(**3): 66-78 (2001)

[30] Adnan el Shukrijumah is a Most Wanted Terrorist

http://www.ianlivingston.com/archived/2004/shukrijumah.htm

[31] Jackson Joab, NSA and Social Networking: TRENDS & TECHNOLOGIES that affect the way government does IT, Government Computer News, 24th July 2006. Available online

http://www.gcn.com/print/25_21/41403-1.html

[32] i2: Analyst's Notebook (Accessed on November 14, 2006.

http://www.i2.co.uk/Products/Analysts_Notebook/default.asp

[33] Wheeler, L. Scott , Dirty Bomb Plot underway in US? (Accessed on November 14, 2006)

http://www.worldnetdaily.com/news/article.asp?ARTICLE_ID=35339

[34] Adnan Gulshair el Shukrijumah (Accessed on November 14, 2006)

http://en.wikipedia.org/wiki/Adnan_Gulshair_el_Shukrijumah

[35] Ramzi Yousef (Accessed on November 14, 2006)

http://en.wikipedia.org/wiki/Ramzi_Yousef

[36] September 11, Attacks (Accessed on November 14, 2006)

http://en.wikipedia.org/wiki/September_11,_2001_attacks