Detecting Key Players in Terrorist Networks

Ala Berzinji

UPPSALA
UNIVERSITET

**Teknisk- naturvetenskaplig fakultet
UTH-enheten**

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
<http://www.teknat.uu.se/student>

Abstract

Detecting Key Players in Terrorist Networks

Ala Berzinji

The interest in analyzing loosely connected and decentralized terrorist networks of global reach has grown during the past decade. Social Network Analysis (SNA) is a mature approach towards understanding terrorist networks since it can be used to analyze a networks structure and properties and to detect important persons and links.

In this work we study decentralized terrorist networks with different types of nodes. The nodes can be either organizations, places or persons. We use a combination of different centrality measures to detect key players in such networks.

Handledare: Lisa Kaati, Ahmed Rezine
Ämnesgranskare: Parosh Abdulla
Examinator: Anders Jansson
IT 11 060
Tryckt av: Reprocentralen ITC

To all who are closest to my soul

Acknowledgements

I am heartily thankful to my supervisors Lisa Kaati and Ahmed Rezine whose support and help enabled me to undertake and successfully complete this thesis work. Their supervision has been truly valuable for me. I learnt a great deal from them. I am extremely grateful to my mentor Prof. Parosh Abdullah whose encouragement landed me in Uppsala University and with whose help and support I won this success.

I would also express my sincere thanks to PhD student Muneeb Khan for constructive discussions and help from the beginning of my studies until the last day. In addition to that I thank all my friends, especially Tara Qadir who has always been there for me, and Hussein Muhammad who helped me with accommodation when I first arrived in Uppsala.

I am grateful to Dr. Anders Berglund and Ivan Christoff for their help on several occasions during my Masters program. Your support is highly appreciated.

And of course I want to thank my family for all the support, love, encouragement and prayers for my success in life. They were always there to help me.

And last but certainly not least, I would like to express my deepest love and gratitude to my Husband. Without his support this project would not have been possible. He helped me with his love and all his valuable information to achieve this goal.

Contents

1	Introduction	3
2	Preliminaries	6
2.1	Graphs and Networks	6
3	Centrality	7
3.1	Degree Centrality	8
3.2	Betweenness Centrality	9
3.3	Closeness Centrality	11
4	Terror and Terrorist Networks	13
4.1	Structure of a Terrorist Social Network	15
4.2	AIntP-3 data model	17
5	Detecting Key Actors in Terrorist Networks	21
6	Algorithm for Discovering Finance Manager	24
7	Case Study	26
7.1	Implementation	26
8	Conclusions and Future work	27

List of Figures

1	Undirected graph	6
2	Degree Centrality.	9
3	Betweenness Centrality.	10
4	Closeness Centrality.	12
5	Alternative Closeness Centrality.	13
6	A decentralized terrorist network.	17
7	A terrorist network.	27

1 Introduction

Social network analysis (SNA) is a set of powerful techniques that can be used to identify clusters, patterns and hidden structures within social networks. Social networks can be found in a number of different areas and SNA can be used to analyze a wide variety of different problems. SNA has many applications in health, political science, economics, sociology etc. A social network is usually represented by a graph, since graphs provide an intuitive idea of how the network is constructed and it is easy to analyze using mathematical methods and software tools. SNA can be used to deduce useful information from a social network, such as:

- analyzing information flow through the network,
- extent of information reach or spread within the network when propagated by a given node,
- identifying specific paths taken by the information to flow from one node to another,
- identifying sub-groups,
- discovering non-obvious relations between actors, and
- identifying nodes that are directly or indirectly connected to most other nodes in the social network.

SNA tools can also be used to identify key actors involved in spreading information. If the nature of the information is identified as illegal and threatening, damage control can be done by stopping the spread of information. This can be done by attacking the nodes acting as key actors or connecting points. The focal point of SNA is the associations and relationships of social entities in social and behavioral sciences. The social network perception has been developed by researchers belonging to various sciences such as sociology, psychology and anthropology.

SNA is a set of techniques that focuses on identifying individuals and comparing the relationships between them. These techniques can categorize individuals into a number of groups based on their attributes (for example friendship, kinship, common interest, financial exchange, dislike, beliefs, knowledge or prestige) in order to model real world interactions within the

network. When the relationships and information of flows become visible, then useful information can be deduced from the graph. For example, techniques from SNA can be applied to actors within groups to identify the persons who have the central roles in the network, who are isolated, or who have the most relations etc. This information can help to improve or even stop information propagation within the network.

The notion of a social network and the techniques of SNA have attracted considerable interest and curiosity from the social and behavioral science community in recent decades. Much of this interest can be attributed to the appealing focus of SNA on relationships among social entities, and on the patterns and implications of these relationships [37]. In SNA, the characteristics of persons or groups are less essential than their associations or the links between the actors inside the network. Social networks have been used extensively to study how relations associate actors, and have also helped in revealing non-obvious (or hidden) relations and information propagation paths between two or more nodes in the network.

With the growing threat of global terror, the interest in efficiently analyzing criminal and terrorist networks of global reach has increased significantly in the past decade. Analyzing criminal networks (sometimes also known as dark networks) and the information flow through them is especially important to establish early warnings about possible future threats emanating from the activities of such groups.

Criminal and terrorist social networks can often be divided into hierarchies with actor roles varying from ordinary operatives to the finance manager and leader of the network. The most important roles include the leader who acts as the guide and mentor for the whole group, besides giving directions about the operations that need to be planned and carried out by the group. The finance manager deals with the execution of planned events and also manages the financing of all other operatives of the group. So the finance manager is usually the node that is directly connected to most other nodes in the network. Other manager roles may include media manager – responsible for propaganda, claiming responsibility for terrorist events and promotion of the group and its objectives on media, military manager – responsible for arranging equipment used in terror related incidents, and operatives – responsible for carrying out the planned events.

In the post-9/11, war on terror era, terrorist organizations around the globe have evolved a decentralized strategy to carry out successful operations in the Middle East [31]. A decentralized approach in practice puts almost the

entire responsibility related to operations on the finance manager. Since the finance manager is the one to give directions for carrying out various operations to other group members in addition to providing them with finances. Therefore, it is extremely important to be able to detect the finance manager from a given structure of a social network in order to effectively counter the operations of the terrorist network.

In this work we present an algorithm that can be used to find the finance manager of a decentralized terrorist network. The network is represented using a subset of the categories that are present in the NATO AIntP-3 data model. The finance manager is the node in a network that is most operationally central, active and that acts as gateway in the network. The finance manager is detected using a combination of different well-known centrality measures.

Related work. Terrorist social networks are also known as dark networks. Some recent studies like Daninig et al. [14] have explored the use of SNA methods to analyze dark networks. His works have mostly focused on assigning roles to actors in the network. In [30] the use of centrality measures to identify key actors in criminal networks is explored and in [36] centrality measures are used to identify the group leader of the September 11th hijackers. However, none of these operates on networks with different categories of nodes.

Xu and Chen use SNA methods in [38] to determine the leader and gatekeeper role for individual nodes in addition to hierarchical clustering methods to identify subgroups within criminal networks. In [20] Memon et al. provide a novel algorithm to automatically detect the hidden hierarchy in terrorist networks. The algorithm is based on centrality measures that are commonly used in social network analysis. Two measures of centrality: termed degree, and Eigenvector centrality are introduced as well as a novel dependence relation.

Memon developed the Investigative Data Mining (IDM) toolkit for subgroup detection and terrorist intelligence analysis [19]. The techniques permit to isolate all those nodes in a social network that are directly or indirectly connected to a specific node. Also, the techniques allow to identify all the paths (in the network) that connect all nodes identified as terrorists belonging to a specific group. Further, shortest paths between all terrorist nodes can also be identified. The techniques make also use of centrality measures

combined with data mining techniques to find the hierarchy of a group and the leader of the group. However, our work is different from theirs because we consider decentralized terrorist networks, whereas their work is mostly related to centralized hierarchy.

2 Preliminaries

2.1 Graphs and Networks

A graph G consists of a pair (V, E) where V is the set of nodes, and E the set of edges that connect the nodes. We assume in the following a graph $G = (V, E)$. Social networks are naturally modeled using graphs where nodes represent actors and edges relations between them. In the following, we often say network to mean the graph modeling it.

Figure 1: Undirected graph

Edges might be undirected or directed depending on whether they reflect symmetrical or not symmetrical relations between actors. We write $uv \in E$, for $u, v \in V$, to mean that there is an edge between nodes u and v in the graph (V, E) . Most SNA techniques focus on undirected graphs since the considered relations are typically mutual and bi-directional. Unless otherwise specified, we only consider in this work techniques for undirected graphs. In other words, we consider that the pairs in the set of edges E are unordered. Also, and otherwise specified, the relations we represent with the edges are not

reflexive, i.e., there are no self loops in E . Figure 1 illustrates an example of an undirected graph.

Basic notions and notations. Many graph algorithms initially select a single node in V and refer to as the *ego node*. This node can be predefined or obtained either randomly or using some calculations. Given a set S , we write $|S|$ to mean the size of S . For instance, we write $|V|$ (respectively $|E|$) to mean the number of nodes (respectively of edges) in the graph. Given a node $u \in V$, all nodes connected to u in E are neighbors of u and make up its *neighborhood*, written $N_G(u)$. In other words, $N_G(u) = \{v \text{ s.t. } uv \in E\}$. The *degree* of a node u , written $d_G(u)$, is the number of nodes in G having an edge with the node u , i.e., $d_G(u) = |N_G(u)|$. Given two nodes u and v in V , a path between u and v is a succession of edges connecting nodes u and v . Formally, a path $\pi(u, v)$ between two nodes u and v is a sequence of edges $w_0w_1, w_1w_2, \dots, w_{n-1}w_n$, with $w_0 = u$ and $w_n = v$. We write $w \in \pi(u, v)$ to mean that the node w participates in the path $\pi(u, v)$, and $|\pi(u, v)|$ to mean the length of the path. The length of a path is the number of occurrences of the edges participating in the path. The *geodesic distance*, written $d_G(u, v)$, is the length of a shortest path connecting the two nodes u and v if such a path exists, and is undefined otherwise. The geodesic distance between a node and each other node in the simple unweighted graphs we consider can be obtained using a breadth first traversal of the graph, i.e., iteratively exploring all neighboring nodes to those nodes that have already been searched. We use $SP_G(u, v)$ to mean the set of shortest paths between nodes u and v in the graph G .

3 Centrality

Analyzing social networks typically aims at categorizing and identifying the roles played by the participants in the network. The analysis estimates the relation of each node to the other nodes. In this context, *centrality* measures are commonly used and aim at capturing the relative importance of nodes in a network. To achieve this, the centrality of a node takes into account how other nodes in the network are related to the node through direct or indirect relations. Intuitively, centrality should give a maximal value to the node in the central point of a star network, and a minimal value to the other nodes. There are various measures of centrality that determine the importance of a

node in the network by taking into account different aspects of the relations between the nodes in the network. The most commonly used centrality measures are degree, eigenvector, betweenness and closeness centrality.

In fact, centrality is also one of the network properties that have been frequently used to study actors or events in terrorist social networks. The most used centrality measures in this context are degree centrality, betweenness centrality, and closeness centrality. We introduce these in the following, together with a particular closeness centrality measure, that we call parameterized closeness.

3.1 Degree Centrality

Degree centrality estimates how important a node is by analyzing the number of direct relationships it has with other nodes in the network. The degree centrality of a node simply corresponds to the degree of the considered node [10]. This measure can be normalized by the number of nodes in the graph G less one (recall we exclude self loops):

$$\mathcal{C}_G^d(v) = \frac{d_G(v)}{|V| - 1}$$

In the case of a directed network, two types of degree centrality are considered: in-degree and out-degree. Whether directed or not, the idea is that the more edges a particular node participates in, the higher is its degree centrality value. This is illustrated in the Figure 2. The actor with the highest degree centrality is considered to be the most strongly (or most frequently) connected node in the network. Such a node holds an advantaged position in the network in terms of connectivity with other nodes which gives it a key role to propagate information. In other words, degree centrality of an ego node is a measure of immediate influence, that is, what proportion of the nodes in the network are influenced by the ego if the latter influences its neighbors with a piece of information and none of the influenced nodes is allowed to further spread the information. The higher the proportion of nodes influenced, higher will be the degree centrality of the ego node.

Other variants of degree centrality (ex [4]) count the number of nodes related by paths of lengths less or equal to some predefined number (one for the original definition). The idea is that a node with few edges related to nodes with many edges still can have a high centrality (degree centrality).

For instance, when the degree centrality is obtained with respect to maximal path lengths of two, it measures how much influence a node has if it influences its neighbors with a piece of information and the influenced nodes in turn are allowed to spread it to their neighbors, but the latter nodes are not allowed to further pass the information. A major advantage of degree centrality is its simplicity as it only takes into account the immediate neighborhood of a node when computing its centrality.

Figure 2: Degree Centrality.

Figure 2 illustrates an example graph where the node with the most edges (node H) has the highest degree centrality according to the first order definition of degree centrality.

3.2 Betweenness Centrality

In *betweenness centrality*, higher betweenness is attributed to nodes that occur most often on the shortest paths between other nodes. Nodes that have a higher probability of being located on the shortest path between other distinct nodes have higher betweenness than other nodes. Such nodes can also be described as gateways. Nodes with high betweenness have a control over

the data flowing among the different groups of nodes in the network, since such nodes often act as bridges. In criminal or terrorist networks, nodes with high betweenness usually indicate the most important or involved actors. An actor with a high degree of betweenness centrality usually holds a favored position in the network. Such a node has greater control over information propagation within the network and represents a bridge which can potentially be a single point of failure. Disconnecting such nodes can effectively disrupt communication within the network.

Figure 3: Betweenness Centrality.

Original definitions of betweenness centrality ([11]) for an ego node v compute all shortest (geodesic) paths among all pairs of nodes in the network. More specifically, the set $\{\pi(s, t) \text{ s.t. } |\pi(s, t)| = d_G(s, t)\}$ is computed for each pair of nodes $s, t \in V$ different from v . Then, the number of times the ego node v was found on these shortest paths, i.e., the size of the set $\{\pi(s, t) \text{ s.t. } v \in \pi(s, t) \text{ and } |\pi(s, t)| = d_G(s, t)\}$ is counted. Betweenness is then calculated as follows:

$$\mathcal{C}_G^b(v) = \sum_{s, t \in V \setminus \{v\}} \frac{|\{\pi(s, t) \text{ s.t. } v \in \pi(s, t) \text{ and } |\pi(s, t)| = d_G(s, t)\}|}{|\{\pi(s, t) \text{ s.t. } |\pi(s, t)| = d_G(s, t)\}|}$$

Calculating all geodesic paths between all pairs is costly. Instead of using geodesic paths, as in the original definition, different alternatives have been

considered. One alternative [5] is to only consider, when computing betweenness centrality for an ego node, the network resulting from the nodes directly related to it together with the edges among them. Other alternatives involve statistical sampling in order to randomly select paths between two random nodes in the network and then to count the number of times the ego node appears on the randomly selected path.

Figure 3 illustrates an example graph pointing out the node with the highest betweenness centrality (node K).

3.3 Closeness Centrality

Closeness centrality of a node in a network is another variant for measuring the centrality of a node. It is a measure of how close a node is to all other nodes in the network (directly or indirectly). Closeness centrality measures how quickly a node can access information through other nodes in the network. A node with a high closeness centrality has a short path to other nodes in the network, and can reach them (i.e. propagate information to them) quickly. Such a node has high visibility as to what is happening in the network, and that is because information in the network may usually flow through nodes with high closeness centrality.

Typically [12], closeness centrality is measured using the geodesic distance (shortest path). The node with the highest closeness centrality is the one with the smallest distance to all other nodes in the network. One way to take into account the distance of a node v in a graph $G(V, E)$ to all other nodes is to sum the distances between v and each of the other nodes:

$$\sum_{u \in V \setminus \{v\}} d_G(v, u)$$

Where $d_G(v, u)$ is the geodesic distance between the nodes v and u . Closeness centrality is then defined as the inverse of the average of this sum:

$$\mathcal{C}_G^c(v) = \frac{|V| - 1}{\sum_{u \in V \setminus \{v\}} d_G(v, u)}$$

Measuring closeness centrality for all nodes boils down to a breadth first search of the entire network for every node.

As a result, closeness centrality does not scale as well as degree centrality. Nevertheless, the obtained result takes into account the network as a whole, instead of limiting to a local, and possibly misleading fraction.

Figure 4 illustrates an example social network pointing out the node with the highest closeness centrality (node J).

Figure 4: Closeness Centrality.

In this work, we came up with an alternative definition of closeness centrality and did check its usefulness by experimenting on a number of benchmarks. In fact, we thought this alternative definition was new till we realized it had already been discovered in [33, 6]. The idea is to sum, when computing closeness centrality for a node v , for each node u different from v , the result of applying a strictly decreasing positive function α to the distance $d_G(u, v)$ between nodes u and v , formally:

$$\sum_{u \in V \setminus \{v\}} \alpha(d_G(v, u))$$

For example, when $\alpha : x \mapsto x^{-1}$ we get the definition in [33], and when $\alpha : x \mapsto 2^{-x}$ we get the one in [6]. Unlike the original definition of closeness, this definition is well suited for disconnected graphs when $\alpha(x)$ takes 0 for infinite x . In the rest of this thesis, we say alternative closeness centrality to mean the closeness centrality obtained by using the function $\alpha : x \mapsto x^{-1}$, and write:

$$\mathcal{C}_G^{ca}(v) = \sum_{u \in V \setminus \{v\}} (d_G(v, u))^{-1}$$

Figure 5 illustrates an example social network pointing out the node with the highest alternative closeness centrality (node H).

Figure 5: Alternative Closeness Centrality.

4 Terror and Terrorist Networks

Terror is a very complex word in the political and security world. Terrorism, like diplomacy, and war, is another political tool to achieve different goals. Because of this, terrorism has a relationship with the social life of people and for the same reason even politics and management events of the society are under the influence of terrorism. Illegal and disruptive activities by terrorist and criminal networks can cause significant loss of life, large economic losses and disastrous social and environmental impacts. When terrorism or extremism has a significant role or influence in a society, then the future of that society will be dependent on the strength and influence of extremist organizations and their views.

There is no good definition to distinguish between terrorist organizations and freedom movements. For example, an incident involving oppositions can be presented by the government as terrorism, whereas the opposition can claim that it is a freedom fighter movement. However any incident that is used to terrorize society and/or disrupt social life can be considered a terrorist event [3]. In a modern world terror can be used as a tool to intimidate groups,

Date	Location	Incident
7/8/98	Nairobi, Kenya	US Embassy bombing in Kenya
7/8/98	Dar es Salam, Tanzania	US Embassy bombing in Tanzania
11/9/01	New York and Washington, US	Attacks on WTC Twin Towers and Pentagon
11/3/04	Madrid, Spain	Attacks on commuter trains
7/7/05	London, UK	Attacks on public trains and buses
25/12/09	Detroit, US	Attempted attack on Detroit bound plane
11/12/10	Stockholm, Sweden	Stockholm suicide bombing
27/2/11	Sulemanyah, Iraq	Attack on American University, Sulemani

Table 1: Some recent terrorist events around the world.

society, people and markets. Terrorist groups can use media, biological, chemical, cyber-space, Internet and ordinary ammunition as their weapons. Table 1 shows a list of some famous terrorist attacks since 1998.

There are several active terrorist organizations in the world. Some of the active terrorist organization operating in the Middle East are:

- Al-Qaeda (Qa'idat al-Jihad, The Base)
- Taliban
- The 1920 Revolution Brigades (Kata'ib Thawrat al-Ishreen)
- Al-Qaeda in Iraq (al-Zarqawi Network, Monotheism and Holy Struggle, Organization of Jihad's Base in the Country of the Two Rivers, Tawhid and Jihad)
- Supporters of Islam (Ansar Al Islam)
- Islamic Army in Iraq (al-Jaish al-Islami fi al-Iraq)

- Tehrik-e Taliban Pakistan
- Army of Jhangvi (Lashker-e-Jhangvi)
- Jund-allah (Tanzeem Jund Allah li-Mujaheedi al-Sunnah fi Iran)
- Al-Qaeda in Arabian Peninsula
- Al-Qaeda in Islamic Maghreb

Terrorist networks A terrorist network consists of a group of people that are members of or somehow related to a terrorist organization. After 9-11 the structure of terrorist networks changed. The structure of the networks became more decentralized and the geography of the network became more global [32]. By analyzing terrorist networks related to well-known terrorist attacks such as: 9/11 [18], the London bombings [8], Detroit Christmas in 2009 [27], the Times Square event in 2010 [26], Stockholm in 2010 [1] and in Sulaimanyah 2011 [9] we can observe that all these terrorist networks are global. The members of the networks did not stay in one place since they needed military training, finances, and indoctrination.

4.1 Structure of a Terrorist Social Network

Terrorist social networks can be represented as graphs where nodes in the network represent actors or groups, and the links between the nodes demonstrate their relationship with each other. The working structure of a terrorist social network was published in 2010 by the counter terrorism agency of Iraq [34]. The report discussed operational and general working structure of terrorist groups in the region. They mentioned that modern terrorist groups are decentralized when they work.

A network that is centralized consists of one or a small number of very central nodes. Such nodes connecting several other nodes in the network indirectly are called Hubs. Such a network may break up into separate dislocated sub-networks, if any of the central nodes is removed. A network that is centralized only around a well connected hub may fail easily if that hub is removed or put out of action. Hubs are nodes with high degree centrality as well as high betweenness centrality (see Section 2 for definitions).

Contrary to that, a network which is less centralized has fewer points of breakdown. It can tolerate several points of failures caused by attacks.

According to Hopkns [13] the challenges in mapping covert networks (criminal networks working in secret) are:

- Incompleteness: the inevitability of missing nodes and links that the investigators will not uncover.
 - Data may be subject to self reported bias, and useful information is not readily available to the public, or at least not easily found.
- Fuzzy boundaries: the difficulty in deciding who to include and who not to include into defined cluster.
 - Data may be biased toward leaders and members captured or identified in an attack.
- Dynamics: These networks are not static and are always changing.

In this work, we focus on finding the most active actors of decentralized terrorist networks. Experience has shown that no terrorist acts can be committed without material supports, and at the same time the cases of Nizal Muhammed (London 7/7 suicide bomber), Umer Farooq (London 7/7 suicide bomber) and Taimour Abdul Wahab (Stockholm suicide bomber) proved that Internet contacts and relations are one of the most important methods to recruit and organize people for terrorist networks.

From the information published in the book [34] we can derive an example of the structure of a decentralized terrorist network. A simplified example of such a network can be found in Figure 4.1. Such a network is led by a leader who mostly acts as a mentor and only provides guidance on how to organize and motivate the group operatives. The actual activities are managed by the finance manager on behalf of the leader. The finance manager acts as the defacto leader of the group. In addition to the finance manager, there are other managers with limited roles such as organizing media propaganda for the group, organizing security related matters for the operatives and managing equipments for militant operations. All group members holding managerial roles are directly supervised by the finance manager. In addition to the managers, the finance manager also provides direct instructions to the operatives directly involved in militant activities.

The finance manager is the one who occupies the most central and active role in a decentralized terrorist network. The finance manager has the

Figure 6: A decentralized terrorist network.

most contact with all other members of the group. Besides that, the finance manager is the only one having direct contact with the actual leader of the group. The leader of the group only sets the goals for the group, and the finance manager manages to achieve those goals.

4.2 AIntP-3 data model

NATO (North Atlantic Treaty Organization) is a collective defense organization based on inter-governmental military alliance of its member states. Being a large inter-governmental defense organization, NATO has requirements for standardized information exchange at all levels, especially in areas of intelligence sharing. These standards are important to avoid confusion in terminology within the organizations. They also ensure efficient, quick and secure transfer of important and sensitive information across relevant or concerned parties within the organization.

The AIntP-3 data model is a standard that is used within NATO to facilitate the electronic and manual exchange of intelligence data between

databases. The AIntP-3 data model defines several categories to capture information about concerned persons, organizations, events, places and equipments. These categories are related to each other with the help of well-defined sub-categories.

In this work we use a subset of the AIntP-3 data model to represent and structure terrorist networks. Using a well-defined standard to represent social networks and to define algorithms is crucial if the methods are to be applicable on information coming from several different sources. For instance, it is common in the Middle East that NATO gathers intelligence on terrorist networks and communicates relevant information to regional law enforcement agencies and other government institutions.

AIntP-3 Categories. There are five main categories in the data model:

Equipment - Materials or tools used to equip investigators or investigative organizations to fulfill their roles. Such equipments or installations may be owned and/or operated by military or civilian and may exist on land, in air space, in outer space (as satellites) or under the sea.

Organization - An organization is a grouping of individuals, with a well-defined hierarchy, aiming to achieve a common goal. Each individual in the group is assigned a role to complete individual tasks aimed towards achieving the goal. An organization can also be of criminal orientation. An organization in general may be divided into sub-organizations with their own specific hierarchy and well-defined sub-functions.

Place - Point or area on earth or space, identifiable through a set of reference coordinates, that is occupied by an entity (unit, equipment, person or organization) or is associated with the occurrence of an incident or event. A place may be a reference to a natural or man made feature and also covers installations and facilities, defined as:

- Installation - a man made feature existing on earth or in space with a well-defined purpose
- Place - besides the reference coordinates, also describes the attributes, the physical characteristics (including description of surroundings and terrain), the current status and condition, and the ability of the location to fulfill any defined interests.

Biography - The description of the appearance of an individual, personal and professional attributes, personal background (family, education, career etc.) and social behavior patterns.

Event - The description of an incident or occurrence that has some intelligence significance. An event can usually be broken down into several smaller sub-events that are in some way immediately or remotely (loosely) related.

Sub-categories Three sub-categories can be used to assist the five main categories of intelligence, enabling more specific intelligence to be captured and expressed. The sub-categories are:

Net-links - A net/link is the physical or virtual connectivity that exists between two nodes in a network. This sub-category allows description of virtual and physical links that can exist between two entities. Physical links can be systems such as pipelines, roads, railways, waterways etc, between one or several Places. Virtual links make use of the electro-magnetic spectrum (EMS) such as television, radio, radar, satellite links and Internet, and may define the links between organizations. Virtual links may also be established by the equipment used by the investigators to connect to other persons or organizations.

Targeting - This sub-category describes the associations between organizations, places, equipments, net-links and targeting information. The targeting data recorded in this sub-category fulfills the following functions

- It accurately identifies and categorizes targets.
- It records the results of post-strike analysis (Battle damage assessment).

Imagery and documents - This sub-category relates imagery/documents to an entity's root record under one or more of the five main categories. It also describes the technical details relating to the imagery document.

Each category consists of two structures (root record structure and relationship structure) within which information or intelligence data is stored according to data groups, data element and data sub-element fields.

Root record Each of the main categories of intelligence is sufficient for recording details of single entities, such as a person, barracks, or a single piece of equipment such as a tank or satellite. The description of each entity consists of a series of individual items of data, describing the entity's qualities or its attributes, such as the height of a person, his date of birth or his personal qualifications. These data items are stored in individual fields set in a structure whose format differs for each category. The aggregate of these data items, which represent a single entity within a category, is known as the root record. Within the structure of each root record, a small number of fields are designated to form the record key. The data contained in these fields is unique to that instance of a root record and the record key may therefore be used as an abbreviated identifier of all the data in the record.

Relationship Structure In many cases there will be a requirement to express information or intelligence that is represented by two entities (which may be related), for example, the type of aircraft stored in a hangar, the unit involved in the engagement or the captain responsible for flying the aircraft. The relationship structure expresses this information by linking one root record to another, and contains fields which both identify the related entities and their relationship. The relationship is described by using a limited number of items of data, such as those setting out the duration of the relationship or the nature of the relationship. In the same way as in the root record, these data items are stored in a series of structures.

Any of the five main categories of intelligence may form a relationship with any other category. The relationships between the five main intelligence, may be one-to-one, many-to-one or many-to-many. With the exception of imagery/documents, other sub-categories (since they have a specialized application) are limited in the number of relationships that they can form with the main categories. The imagery/documents and targeting sub-categories do not have separate root records and relationship structures. The record key contains the fields necessary for them to link to other entity records.

5 Detecting Key Actors in Terrorist Networks

Due to security reasons, the leaders of terrorist networks limit their relations (or interactions) with other members of the terrorist network. The member who manages the finances of the terrorist network has most frequent contacts with the leader of the terrorist network. In most cases, this member acts as the deputy or alternative leader of the network due to the following reasons:

- The member has most contacts with all other operatives in the network.
- The financial resources for the entire group are controlled by the member, and the leader of the network relies on this member for all operations of the network.

Finances are an important part of a terrorist operation. After studying several terrorist operations, Levitt states in [17] that it is clear that without funds no terrorist events can take place. Also, Grimes proposes in [7] that tracking financial circulation in a network is a way to discover every member in the network.

The finance manager is a key player in a terrorist network. The finance manager holds information of most of the members and the activities carried out by the network since it is not possible to carry out any terrorist related activities without material support (which requires finances).

The US Department of the Treasury mention in [25] that after the terrorist attacks on September 11, 2001, they have initiated the Terrorist Finance Tracking Program (TFTP) to identify, track, and pursue terrorists in networks such as Al-Qaida. The Department of the Treasury is uniquely positioned to track terrorist fund flows and assist in broader governmental efforts to uncover terrorist cells and map terrorist networks within the USA and around the world.

Because of the risk of being caught by security agencies, the information about the leader is usually very hard to get, so it is easier to track or find out about the finance manager. The finance manager is second in command in decentralized networks.

Terrorist organizations in Middle East almost never have any female operatives, though there have been a few recent exceptions to this trend, but only limited to ground operatives. Therefore, we assume the finance manager to always be male. Since the finance manager has most of the operational

capacity in a terrorist network tracking his activities or apprehending him will cause the most harm to the operational capabilities of the group.

In [16] the tracking of two identified terrorists is described. They found that the finance manager acted as a local leader, and they could track his relations and find the actual leader through him.

Only the leader can make a decision about an operation and all operations need financial resources. Therefore, the leader depends on the finance manager. This means that usually the leaders' only relationship in the network is with the finance manager, which in turn is in direct contact with other involved (or required) actors and operatives [34].

In post 9/11 the structure of terrorist networks are decentralized and it is a challenge to track the most important actor within terrorist networks. When terrorist organizations are formed there are always some kind of hierarchy. However, due to security reasons the actors are defensive and mostly on the run, so they will not appear in hierarchical structure [35, 24, 29].

Lately, the terrorist network Al-Qaeda which is believed to be most strongly present in Pakistan, Iraq, Yemen, and the Magreb (Morocco to Libya) has started to become extremely decentralized in its operations. In [20, 22, 21] Memon mentions that Al-Qaeda works as a decentralized network.

In [2], Anderson has worked to determine the disruption of terrorist financing as an approach to effectively disrupt terrorist operations. He discovers that it is possible to disrupt financial flows.

In this work, we present a method for detecting the finance manager in a decentralized network. This will effectively result in finding out the members of the group and the new pattern of the network. The finance manager has to fund the activities of terrorist groups in different geographic locations and has to provide funds to the operatives performing those activities. For this reason the finance manager has the most relations in many different regions. We can find the node in the network with the most relations to different places with the measure of degree centrality [15]. This is due to two reasons, the finance manager has cluster relations to almost all other members of the group.

- The salary of the group members is managed by the finance manager (in case of Al-Qaeda in Iraq, all members receive at least 100 dollars every month from the finance manager [34])
- Because of the operations of the group, the finance manager supports

every active operative with extra funds and therefore has close relations with them [34].

One of the problems of the terrorist groups is the procurement of funds and their secret transfer to other people in other places. After 9/11 government security organizations around the world have focused on transfers of funds when they are large in amount, and also when funds are transferred from countries with conflicts to EU or North America [28]. The finance manager of a group needs to avoid being tracked by any security agency and avoid tracking by:

- Acting as the supervisor for his organization. He organizes different projects and organizations globally and locally overseeing different businesses. This helps procure/generate funding for various activities of his group [34].
- The finance manager has relations besides the organizations that he oversees with other groups or organizations that provide funding to his group [34].

From this we can observe that the finance manager acts as a gatekeeper between the organizations that provide funds to his organizations. Besides that he also manages the finances of his group's organizations and businesses. As a result, the finance manager is represented by a node with high betweenness centrality with local and global organizations within the network [23].

We are interested in developing a technique to track the finance manager in terrorist organizations that follow the modern trends of post 9/11 decentralized terrorist networks. So finding the finance manager will require employing a combination of strong centrality measurements as

- Finding the node with most relations with other nodes, using degree centrality.
- Finding the node that has the most relations with other places, using degree centrality.
- Finding the node with the most closeness to all other nodes, both with the original and the alternative definitions of closeness centrality.
- Finding the node that is the gateway between all the organizations that have relations with this group using betweenness centrality.

Following these steps can help in narrowing down the focus to the key actors in the social network and finding out the finance manager.

6 Algorithm for Discovering Finance Manager

In this section we present an algorithm designed to find the node in the network that

1. represents a person having highest degree centrality among all other persons and places,
2. is closest to all other nodes in the network, and
3. has highest betweenness centrality between the organizations.

This way we can discover the node that is operationally most central, active, gateway and controllable in the network.

The social network is represented by an undirected graph (V, E) with n vertices composed of the following subgraphs.

- (V_p, E_p) – graph representing persons as nodes and relations between them. The number of vertices is denoted n_p .
- (V_o, E_o) – graph representing organizations as nodes and the relations between them. The number of vertices is denoted n_o .
- (V_{pl}, E_{pl}) – graph representing places as nodes and the relations between them. The number of vertices is denoted n_{pl} .

The sum of all the nodes in the above subsets equals the total number of nodes in the entire graph, i.e., $n = n_p + n_o + n_{pl}$. In addition to the given edges in (V_o, E_o) , (V_{pl}, E_{pl}) , there are three additional subsets of edges that are a part of the set E in (V, E) . These sets of edges are:

- E_{p-o} – edges representing relations between persons and organizations.
- E_{p-pl} – edges representing relations between persons and places.
- E_{o-pl} – edges representing relations between organizations and places.

The purpose of the algorithm is to find a node in the network that represents a person that is operationally most central to the network and also closest to other nodes. The algorithm consists of the following five steps.

1. **Degree Centrality** is calculated for each node v_p in graph (V_p, E_p) ; that is, $\mathcal{C}_{(V_p, E_p)}^d(v_p) = \frac{d_{(V_p, E_p)}(v_p)}{n_p - 1}$ is calculated for each v_p in the set V_p .
2. **Degree Centrality** is calculated for each v_p in V_p in the subgraph $(V_p \cup V_{pl}, E_{p,pl})$; that is, $\mathcal{C}_{(V_p \cup V_{pl}, E_{p,pl})}^d(v_{pl}) = \frac{d_{(V_p \cup V_{pl}, E_{p,pl})}(v_{pl})}{n_{pl} - 1}$ is calculated for each v_p in the set V_p .
3. **Closeness Centrality** is calculated for each node v_p in V_p and with respect to all nodes in (V, E) ¹; that is, $\mathcal{C}_{(V, E)}^c(v_p) = \frac{|V| - 1}{\sum_{u \in V \setminus \{v_p\}} d_{(V, E)}(v_p, u)}$ is calculated for each v_p in V .
4. **Alternative Closeness Centrality** is calculated for each node v_p in the set V_p ; that is, $\mathcal{C}_{(V, E)}^{ca}(v_p) = \sum_{u \in V \setminus \{v_p\}} \left(d_{(V, E)}(v_p, u) \right)^{-1}$ is calculated for each node v_p in V_p .
5. **Betweenness Centrality**² is calculated for each node v_p in V_p and with respect to all organizations in V_o ; that is, $\mathcal{C}_{(V_p \cup V_o, E_{p,o})}^b(v_p)$, defined as $\sum_{s, t \in V_o} \frac{|\{ \pi(s, t) \text{ s.t. } v_p \in \pi(s, t) \text{ and } |\pi(s, t)| = d_{(V_p \cup V_o, E_{p,o})}(s, t) \}|}{|\{ \pi(s, t) \text{ s.t. } |\pi(s, t)| = d_{(V_p \cup V_o, E_{p,o})}(s, t) \}|}$ is calculated for each node v_p in V_p .

Once all the centrality scores have been measured as described in above items, the key actors in the social network can be identified. The nodes with high scores for most of the centrality measures represent all the key actors in the network. The node with the highest score on all the centrality measures represents the finance manager. This is also expected since the finance manager in decentralized networks has all the properties to have largest scores

¹Measuring closeness centrality separately for organizations and places does not produce very meaningful results. We consider all nodes (including persons, organizations and places) and calculate closeness centrality for them, but for targeting purposes the most useful node is the one with the highest Closeness Centrality score and represents a person.

²We only consider nodes and edges between persons and organizations in this case. We do so by measuring betweenness centrality for all nodes between vertices that represent organizations. Our interest here is the node with the highest betweenness centrality score that represents a person.

on all mentioned centrality measures. The most important properties of the finance manager include - (1) most direct or 1-hop neighboring nodes, which makes it score higher on closeness centralities, and (2) most common occurrences on all shortest paths between all nodes in the network, making it score higher on betweenness centrality.

7 Case Study

In this section, we consider a terrorist social network and apply the technique we developed to locate the finance manager in the terrorist social network described in the case study. On February 27 in 2011 - 5 terrorists were killed in Sulaymaniya, Iraq. The terrorists were possibly planning an attack on the American university campus in Sulaymaniya. By using open source data such as public information from newspapers on the Internet a terrorist network was created. The network is presented using some of the categories of AIntP-3 data model. The network is shown in Figure 7.

In the network in Figure 7, the person who assigned the 5 terrorists is called S. S is the brother-in-law of a person called J and he is also acting as Js finance manager. As can be seen in the network, S was involved in several businesses in several countries.

Detecting the finance manager. The possible terror attack was revealed when the finance manager was found by the counter terrorism agency of Iraq. Applying our algorithm on the terrorist network extracted from the planned terror attack (Figure 7) we notice that the actor named S always has the highest score on all centralities between persons, places, organizations and relative to all nodes in the network. S is therefore a possible finance manager of the group since he maintains most of the connections with other members of the group and most connections with places and organization nodes in the network.

7.1 Implementation

We implemented our approach in Java using JUNG (Java Universal Network/Graph Framework). JUNG is a software library that can be used for modeling, analysis, and visualization of data that can be represented as a graph or network. In the implementation an adjacency matrix representing

Figure 7: A terrorist network.

the relationships between all the nodes in the graph is used. Our experiments on the derived terrorist network of Figure 7 confirmed that our approach is able to identify the key actors in the network, and to isolate the finance manager.

8 Conclusions and Future work

In this work we present an algorithm that can be used to detecting the finance manager in a decentralized terrorist network. The finance manager plays a central role and is the most active actor in a terrorist network.

The networks that we consider contains different categories of nodes. The categories we use are a subset of categories that are presented in the NATO AIntP-3 data model. Detecting the finance manager is done by using a combination of different well-known centrality measures on the different categories of nodes in the network.

One direction for future work is to investigate more complex terrorist networks with different relations as well as different types of nodes. To analyze such networks properly new measures and algorithms are needed. Analyzing more complex networks may provide analysts with more information regarding the key players, structure and information flow in the network than using the traditional social networks where only persons and relations are present.

References

- [1] Aswat AllIraq. 5 "terrorists" killed in clashes. 2011. Online article by Aswat Al Iraq: First Independent News Agency in Iraq. Last visited: 12th August 2011.
- [2] M.Wesley J. Anderson. Disrupting Threat Finances: Utilization of Financial Information to Disrupt Terrorist Organizations in the Twenty-First Century. 2007. U.S. Army, School of Advanced Military Studies, United States Army Command and General Staff College Fort Leavenworth, Kansas.
- [3] J. Bolt. Terrorists or Freedom Fighters: What is the difference? 2001.
- [4] P. Bonacich. Power and Centrality: A Family of Measures. *American Journal of Sociology*, 92(5):1170–1182, 1987.
- [5] S. Borgatti. Centrality and network flow. *Social Networks*, 27(1):55–71, Jan 2005.
- [6] C. Dangelchev. Residual closeness in networks. *Physica A: Statistical Mechanics and its Applications*, 365(2):556–564, 2006.
- [7] Lieutenant Colonel Darryle J. Grimes. The Financial war on terrorism, grading U.S. strategy for combining the financing of terrorism. *United States Air Force*, 2006. Joint forces staff collage joint advanced warfighting school.

- [8] R. Esposito and B. Ross. Northwest Bomb Plot Planned by al Qaeda in Yemen. 2009. Online article by abc news. Last visited: 12th August 2011.
- [9] R. Floyd. 'Safe' kurd city is shaken. 2011. Online article by The Augusta Chronicle. Last visited: 12th August 2011.
- [10] Linton C. Freeman. Centered Graphs and the Structure of Ego Networks. 1982. University of California.
- [11] R. A. Hanneman and M. Riddle. Introduction to Social Network Methods. 2005. University of California. Last visited: 1 July 2011.
- [12] F. Hildorsson. Scalable Solutions for Social Network Analysis. 2009. Uppsala University.
- [13] A. Hopkins. Graph Theory, Social Networks and Counter Terrorism. 2010. University of Massachusetts Dartmouth.
- [14] Daning Hu Jie Xu and Hsingchun Chen. Dynamics of Terrorist Network: Understanding the Survival Mechanisms of Global Salafi Jihad. *Journal of Homeland Security and Emergency Management*, 6, 2009.
- [15] E. Kaplan. Tracking Down Terrorist Financing. 2006. Online article by Council on Foreign Relations. Last visited: 12th August 2011.
- [16] V. Krebs. Connecting the Dots Tracking Two Identified Terrorists. 2008. Online article by Orgnet. Last visited: 12th August 2011.
- [17] M. Levitt. Checkbook jihad. 2011. Online article by Foreign Policy. Last visited: 12th August 2011.
- [18] Sujoyini Mandal. Financial Investigation and CounterTerrorism Case Study: 7 July 2005,London. 2005. ICPYTR, RSIS Singapore.
- [19] Nasrullah Memon, David L. Hicks, and Henrik Legind Larsen. How Investigative Data Mining Can Help Intelligence Agencies to Discover Dependence of Nodes in Terrorist Networks. In *Proceedings of the 3rd international conference on Advanced Data Mining and Applications, ADMA '07*, pages 430–441, Berlin, Heidelberg, 2007. Springer-Verlag.

- [20] Nasrullah Memon, Henrik Legind Larsen, David L. Hicks, and Nicholas Harkiolakis. Detecting Hidden Hierarchy in Terrorist Networks: Some Case Studies. *Springer-Verlag*, pages 477–489, 2008. Proceedings of the IEEE ISI 2008 PAISI, PACCF, and SOCO international workshops on Intelligence and Security Informatics.
- [21] Nasrullah Memon, Abdul Rasool Qureshi, Uffe Kock Wiil, and David L. Hicks. Novel Algorithms for Subgroup Detection in Terrorist Networks. *Availability, Reliability and Security, International Conference on*, 0:572–577, 2009. IEEE Computer Society.
- [22] Nasrullah Memon, Uffe Kock Wiil, and Pir Abdul Rasool Qureshi. Practical algorithms for subgroup detection in covert networks. *International Journal of Business Intelligence and Data Mining*, 5:134–155, January 2010. Inderscience Publishers.
- [23] Aliraq news staff 1. Al Qaeda in Iraq turns to mafia-style business, extorting money to fund its activities- report. 2011. Online article by AssafirPress. Last visited: 13th August 2011.
- [24] Deputy Chief of Staff for Intelligence. A Military Guide to Terrorism in the Twenty-First Century. 2005. US Army Training and Doctrine Command, Intelligence Support Activity - Threats. Fort Leavenworth, Kansas.
- [25] Department of the Treasury. Terrorist Finance Tracking Program (tftp). 2011. Online article by U.S. Department of the Treasury. Last visited: 13th August 2011.
- [26] Daily Mail Reporter. British police given more time to question man suspected of aiding botched Stockholm suicide bomb attack. 2011. Online article by dailymail. Last visited: 13th August 2011.
- [27] R. Rivera. A Dread Revived: Terror in the Trunk. 2010. Online article by NYtimes. Last visited: 13th August 2011.
- [28] Terrorist Financing Operations Section. Terrorist Financing. 2007. Online article by FBI. Last visited: 12th August 2011.

- [29] Jacon N. Shapiro and David A. Siegel. Underfunding in Terrorist Organizations. *International Studies Quarterly*, 51:405–429, 2007. Stanford University and Florida State University.
- [30] M. K. Sparrow. The application of Network Analysis to criminal intelligence: An assessment of the prospects. 1991. *Social Networks*. Elsevier 13, 251-274, 3.
- [31] National Commission On Terrorist Attack Upon The United States. The Global Salafi Jihad. 2003. Third public hearing of the National Commission on Terrorist Attacks Upon the United States, Statement of Marc Sageman to the National Commission on Terrorist Attacks Upon the United States by UNT libraries. Last visited: 1 July 2011.
- [32] National Commission On Terrorist Attack Upon The United States. Final Report of the National Commission on Terrorist Attacks Upon the United States. 2004. The 9/11 Commission report executive summary by UNT libraries. Last visited: 1 July 2011.
- [33] F. Agneessens T. Opsahl and J. Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.
- [34] Anti terrorism Agency-Kurdistan Iraq. *The Geosecurity of Mosel*. 2010. PUK media.
- [35] Terrorism-research. Terrorist Groups. 2011. Online article by Terrorism Research. Last visited: 13th August 2011.
- [36] V.E.Krebs. Mapping networks of terrorist cells. 2002. *Connections* 24(3): 43-52 International Network for Social Network Analysis.
- [37] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Addison-Wesley, 1994.
- [38] Jennifer Xu and Hsinchun Chen. Criminal network analysis and visualization. *Commun. Association for Computing Machinery*, 48:100–107, June 2005.