8-2-2011

On the detection of hidden terrorist cells immersed in peer to peer networks

Belinda A. Chiera
University of South Australia

ON THE DETECTION OF HIDDEN TERRORIST CELLS IMMERSED IN PEER-TO-PEER NETWORKS

Belinda A. Chiera
School of Mathematics and Statistics,
University of South Australia, Australia

belinda.chiera@unisa.edu.au

Abstract

Hidden terrorist cells in high dimensional communications networks arise when terrorists camouflage connectivity to appear randomly connected to the background network. We investigate hidden network detectability when the background network does not support terrorist activities. Using two September 11 terrorist networks as the test bed and a network measure called assortativity, we suggest hidden terrorist networks can behave as Peer-to-Peer networks. We compare the September 11 hidden networks with Peer-to-Peer networks containing embedded terrorist networks, as well as with generic Peer-to-Peer networks. Using Peer-to-Peer characteristics and social network group-based centralities, we show that for certain Peer-to-Peer networks it is possible to detect hidden terrorist networks in cyberspace, with potential future application to Instant Messaging and Skype networks.

Keywords

Detecting terrorist cells, Hidden Networks, Peer-to-Peer Networks, Group-Based Centralities, Assortativity

INTRODUCTION

Terrorist attacks are a recurring threat to local livelihood and global wellbeing and in the past decade have been enacted as the large-scale tragedies of September 11, 2001 (United States), October 12, 2002 (Indonesia) and July 7, 2005 (United Kingdom). More recently, the Lockerbie-styled plane bombing attempts (October 2010) reflect the ongoing global threat of terrorism. Concerns have been raised over the growing strength of terrorist organisations such as al Qaeda, which in the last decade has increased both its membership and geographic reach (Farrall, 2011).

A considerable body of research on modelling terrorist networks has developed in recent years. The cornerstone publication of (Krebs, 2002) used social network measures to capture pertinent characteristics of the individuals involved in the September 11 hijackings. In (Hussain, 2010, Borgatti, 2003) key individuals within terrorist networks were identified such that their removal would cause maximal disruption to the network. In (Chiera, 2010) a paradigm shift from the social network analysis of individual terrorists to *groups* of terrorists was introduced, by treating a terrorist cell as a *hidden network* – that is a network of individuals appearing randomly connected to the background communications network whilst preserving connectivity between themselves, to facilitate unimpeded information flow. Using the September 11 terrorist network of (Krebs, 2002), hidden terrorist networks were characterised using group-based social network measures and were shown to be highly visible and therefore detectable.

The conclusions drawn in (Chiera, 2010) however, were for hidden networks immersed in a larger terrorist network. It needs to be conjectured that it may not be sufficient to consider only the social network measure and size of the surrounding network; network *type* may also play an important role in detection. Specifically, *are hidden terrorist networks more difficult to detect when located in a network that does not support terrorist activities?*

In this paper we take a first step towards addressing this question. The ability to distinguish hidden terrorist networks from a larger network backdrop has long been considered vital for successful detection and thwarting a planned terrorist attack (Baumes, 2008), however while previous work has focused on detection dependent upon the background network size, it is not yet sufficiently clear if hidden terrorist network detection also depends on the whether the background network supports the terrorists' online behaviour.

Using assortativity, a network measure of association, we will show that hidden terrorist networks can behave as Peer-to-Peer (P2P) networks. Correspondingly, we generate embedded P2P networks containing hidden September 11 terrorist networks, as well as generic P2P networks, as our test bed. Using the P2P network characteristics degree distribution, clustering and average path length, as well as group-based eigenvector, information, and subgraph centrality, known to be relevant to P2P networks (Estrada and Rodríguez-Velázquez,

2005), we will investigate whether hidden networks are easily distinguishable from non-terrorist P2P network backgrounds.

Next we review the group-based social network metrics used in this analysis. We then determine the assortativity of hidden September 11 terrorist networks (Krebs, 2002), introduce the embedded and simulated P2P networks and investigate their P2P characteristics. Following, we conduct group-based social network analyses of hidden terrorist networks located within P2P networks before giving our conclusions.

CHARACTERISING HIDDEN NETWORK TYPE AND BEHAVIOUR

We wish to characterise two aspects of a hidden network: *type,* used to determine the network category (Newman, 2010); and *behaviour*, designed to capture different facets of communication between individuals or groups. Accordingly, we view a communications network as a graph $\mathcal{G}$ consisting of individuals (nodes) $v_i \in \mathcal{V}$, and edges $e_{ij} \in \mathcal{E}$, which can be viewed as lines of communication connecting individuals v_i, v_j, such that $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$.

A hidden network $\mathcal{H}$ is defined as a sub-network of $\mathcal{G}$ such that $\mathcal{V}_\mathcal{H} \subseteq \mathcal{V}$, $\mathcal{E}_\mathcal{H} \subseteq \mathcal{E}$, with $\mathcal{H}$ treated as a single node, $v_\mathcal{H}$. An example of $v_\mathcal{H}$ is depicted in Figure 1 in which the hidden network members are denoted by green circles. Previous analyses in this area (for example Hussain, 2010, Baumes, 2008, Borgatti, 2003, Krebs, 2002) have focused on detecting the individuals within a hidden network. However here, as in (Chiera 2011, Chiera 2010), a hidden network will be considered as a single network node for the analysis that follows. Thus, for the example network presented in Figure 1, this implies that the network would be considered treated as consisting of nine nodes.

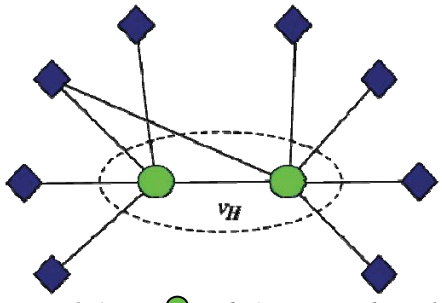

Figure 1: An example hidden network (green nodes) immersed in a larger network (blue nodes). The dashed lines (- -) indicate the boundary of the hidden network node v_H.

The fundamental centrality measure underlying much of the characterisations presented here is degree centrality and while not explicitly analysed in this work, is introduced to facilitate understanding. *Degree centrality* is a measure of local influence in the network (Freeman, 1979) such that a node with high degree centrality would have many direct connections to other network nodes and would thus be considered influential.

For an individual node v_i, degree centrality is the number of edges directly connected to v_i. For a hidden network $\mathcal{H}$, treated as a single node $v_\mathcal{H}$, the normalised group degree centrality C_D is defined as the ratio of the number of non-hidden network nodes connected to $\mathcal{H}$, to the number of non hidden network nodes (Everett and Borgatti, 2005), *viz.*

$$C_D = \frac{|N(\mathcal{H})|}{|\mathcal{V}| - |\mathcal{H}|} \tag{1}$$

where $|\cdot|$ is cardinality, $N(\mathcal{H})$ is the set of all nodes $v_i \in \mathcal{V}$ such that $v_i \notin \mathcal{H}$ but is connected to a member of $\mathcal{H}$, with multiple ties counted once so as not to overinflate the influence of $\mathcal{H}$ in the network. For example, in Figure 1 the normalised group degree centrality of $\mathcal{H}$ is $C_D = 8/(10-2) = 1.00$.

Hidden Network Type: Assortativity

An indication of the type and robustness of a network can be obtained from the *degree assortativity* of a network. Assortativity is the tendency observed in complex networks for nodes to connect mostly with other, similar nodes (Newman, 2010). Similarity is typically defined on the basis of degree centrality, such that nodes with high degree centrality may prefer to connect to other nodes with high degree centrality (*assortative mixing*). In this instance the network is robust in that many network nodes would need to be removed to destroy the overall network structure. Alternatively, nodes with high degree centrality may prefer to connect to nodes of low degree centrality (*disassortative mixing*) making the robustness of the overall network structure more sensitive to node removal.

With a slight abuse of notation, denoting the degrees of nodes $v_i, v_j \in \mathcal{H}$ as k_i, k_j respectively, network degree assortativity is defined as (Newman, 2010)

$$r = \frac{|\mathcal{E}|^{-1}\sum_{e_{ij}\in\mathcal{E}} k_i k_j - \left[|\mathcal{E}|^{-1}\sum_{e_{ij}\in\mathcal{E}} \frac{1}{2}(k_i + k_j)\right]^2}{|\mathcal{E}|^{-1}\sum_{e_{ij}\in\mathcal{E}} \frac{1}{2}(k_i^2 + k_j^2) - \left[|\mathcal{E}|^{-1}\sum_{e_{ij}\in\mathcal{E}} \frac{1}{2}(k_i + k_j)\right]^2}$$

where $|\mathcal{E}|$ gives the total number of edges and the degree centrality of the hidden network $\mathcal{H}$ is defined as in (1). A network is assortative if $r>0$, non-assortative if $r=0$ and disassortative if $r<0$, with $r \in [-1,1]$.

Hidden Network Behaviour: Group-Based Centrality

Eigenvector Centrality

Eigenvector centrality extends the concept of degree centrality to capture the *importance* of a node's connections. Whereas in degree centrality only the number of node connections was of interest, eigenvector centrality takes the further step of determining the quality of these connections — labelling a node of high degree centrality as important only if it is connected to other nodes of high degree centrality. Conversely, a node with high degree centrality connected to nodes with low degree centrality would be classified as unimportant. In this way, eigenvector centrality provides an indication of the most globally central network node (Newman, 2010).

For a network adjacency matrix **A,** with elements $A_{i,j} = 1$ if nodes v_i, v_j are connected and $A_{i,j} = 0$ otherwise, the eigenvector centrality x_i of node v_i is proportional to the average of the centralities of v_i's neighbours

$$x_i = \frac{1}{\lambda}\sum_{j\in M(i)} A_{i,j} x_j$$

where $\mathcal{M}(i)$ is the set of nodes connected to the v_1^{th} node and λ is a constant capturing the relative importance of all nodes in the network. A large value of x_i indicates that node v_i is considered important.

Subgraph Centrality

Subgraph centrality is also designed to extend the concept of degree centrality by taking into account the influence of node v_i beyond its immediate neighbours, to depict the ease with which information is shared throughout the network. Using the concept of a *closed walk*, that is a network path comprised of repeated edges beginning and ending at node v_i, subgraph centrality is a weighted sum of the number of such closed walks of different lengths in the network. Smaller subgraphs (closed walks of shorter length) are given more weight than larger subgraphs, reflecting the immediacy of the influence of node v_i. Subgraph centrality $C_s(i)$ is thus defined as (Estrada and Rodríguez-Velázquez, 2005)

$$C_s(i) = \sum_{k=0}^{\infty} \frac{\mu_k(i)}{k!}$$

where $\mu_k(i)$ is the number of closed walks of length k beginning and ending at node v_i and ! indicates a factorial. Dividing by $k!$ guarantees convergence of the infinite sum, as well as an appropriate weighting, based on walk length. In practice, the summation in $C_s(i)$ is evaluated up to $|\mathcal{V}|$, the number of nodes in the network. Note that the larger the value of $C_s(i)$, the more easily information flows around the network, when originating at node v_i.

Information Centrality

Information centrality is the only measure considered here not based on degree centrality. Rather, it assesses all paths between two nodes to provide an alternative quantification of node importance than that produced by eigenvector centrality, by considering the use of circuitous paths through the network. Circuitous path usage precisely captures the situation of information sharing between terrorists via trusted parties external to the hidden network $\mathcal{H}$ (Newman, 2010, Wasserman, 1994).

Following (Wu et al., 2010) we define information centrality as a function of network efficiency. Efficiency relates the importance of a network node to a decline in communication efficiency if that node is removed from the network. For $d_{vi,vj}$, the length of the shortest path between nodes v_i, v_j, we define the efficiency $E(\mathcal{G})$ of network $\mathcal{G}$

$$E(\mathcal{G}) = \frac{\sum_{v_i \neq v_j \in \mathcal{G}} 1/d_{v_i,v_j}}{|\mathcal{V}|(|\mathcal{V}|-1)}$$

from which a group-based information centrality measure C_I is (Wu et al., 2010)

$$C_I = \left(E[\mathcal{G}] - E[\mathcal{G}_i']\right)/E[\mathcal{G}]$$

where $\mathcal{G}_i'$ is network $\mathcal{G}$ with node v_i removed. The measure is designed such that positive information centrality suggests network efficiency whereas a negative value would be indicative of network inefficiency.

HIDDEN NETWORK ANALYSIS: TYPE

We utilised, as a starting point for the test bed, two versions of the September 11 (9/11) network (Krebs, 2002). The first is the core 9/11 network $\mathcal{G}_C = \{|\mathcal{V}|, |\mathcal{E}|\} = \{19, 27\}$, consisting solely of the individuals who physically hijacked the flights (all nodes except the grey △ nodes in Figure 2). The second is the full network in Figure 2 with $\mathcal{G}_E = \{|\mathcal{V}|, |\mathcal{E}|\} = \{37, 82\}$, containing all individuals deemed complicit in aiding the September 11 terrorist attack. We consider two networks here as a means of capturing differences in hidden terrorist network detectability in different sized networks.

We identified four hidden networks in Figure 2 based on the hijacked flights:

1. **AA Flight # 77**: orange▽ nodes;
2. **UA Flight #93**: blue◆ nodes;
3. **AA Flight #175**: purple■ nodes; and
4. **UA Flight #11**: green● nodes.

A fifth hidden network, **Trusted Priors** (Figure 3), was identified in (Krebs, 2002) as a network of previously acquainted individuals, all of whom believed to have been key individuals in planning the September 11 attack.

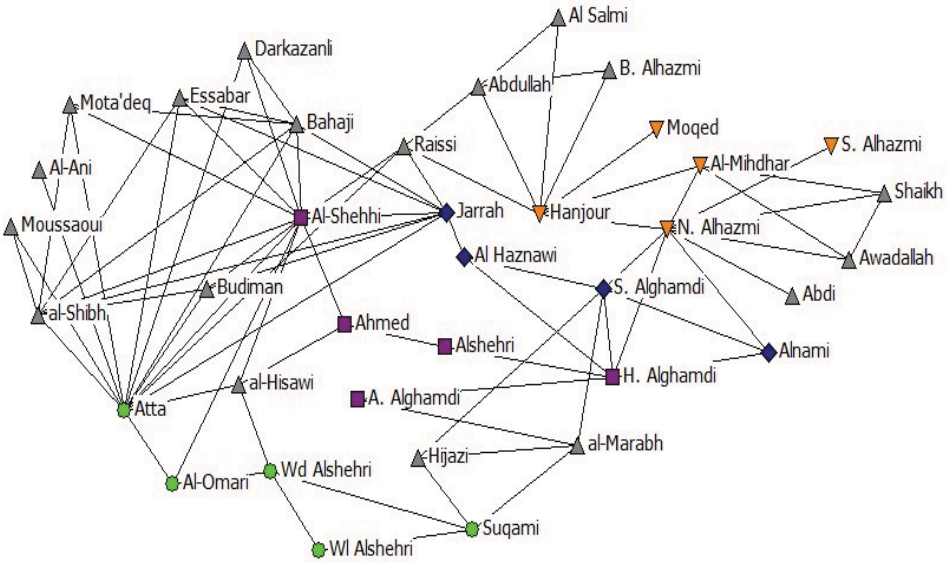

Figure 2: A reproduction of the September 11 Network (Krebs, 2002)

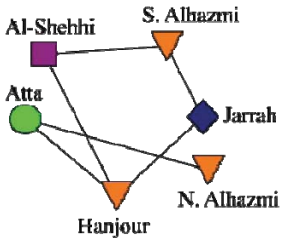

Figure 3: The September 11 Trusted Priors Network identified in (Krebs, 2002)

Table 1: Assortativity (r) for the core and extended 9/11 networks G_C, G_E.

Network Type	#77	#93	#175	#11	Trusted Priors
9/11 Network G_C	0.14	0.00	-0.03	-0.09	-0.14
9/11 Network G_E	-0.06	-0.50	-0.39	-0.07	-0.03

We computed the degree assortativity for both 9/11 networks (Table 1). For the G_C networks only AA Flight #77 was assortative ($r > 0$), seemingly reflecting the fact that the hidden networks for Flight #77 and Flight # 93 are structurally similar. This is not an inconceivable result since previous work (Chiera, 2010; Lindelauf, 2009) indicated a propensity for similar network structures based on size, coupled with a trade-off between the need for information sharing and secrecy.

From Table 1 we see that all hidden G_E networks were disassortative (*r<0)*, particularly for Flight Networks #93 and #175 ($r = -0.50$, $r = -0.39$, respectively). Comparing the internal structures of these two networks with that of their direct neighbours (Figure 2), we see this disassortativity captures the difference between the simple network structures (Flights #93, #175) as opposed to the more complex connectivity of their neighbours (Flights #11, #77).

Comparing the G_E assortativities with those of 27 different network types (Newman, 2010), it was found that many values reported in Table 1 were categorised as Technological, where, for example, $r = -0.39$ for Flight #175 network is close to $r = -0.37$ (the assortativity value for an undirected peer-to-peer (P2P) network, (Newman, 2010)). Comparison of a terrorist network to a P2P network is not necessarily unreasonable since P2P networks have the goals of decentralisation, immediate connectivity and anonymity; characteristics attractive to a terrorist network. Moreover, P2P networks are self-organising such that the failure of a smaller network component is less likely to spread throughout the network; correspondingly, should one hidden terrorist network be compromised, the remaining terrorist cells would remain undetected and able to carry out their mission.

Since our aim is to detect hidden terrorist networks in high dimensional communications networks, we are primarily interested in terrorist communication captured at the Internet routing level, the domain of P2P networks. Combined with the potential for hidden terrorist networks to behave as undirected P2P networks, we wish to determine whether a hidden terrorist network embedded in a P2P network can be distinguished from a generic P2P network structure.

Ideally, a P2P network should satisfy the following properties (Newman, 2010, Wang et al., 2006):

1. It is *scale-free* with a power law degree distribution $p_k = Ck^{-\alpha}$, where p_k is the probability a randomly chosen node has degree k, and C is a constant. The power law exponent α is typically $2 \leq \alpha \leq 3$, capturing the phenomenon that many network nodes have low degree with only a small number having high degree;
2. It is *small-world*, that is the average hop distance between nodes is short and scales logarithmically; and
3. It should have a high *clustering coefficient,* indicating a propensity for two nodes to cluster together with a third node. The global clustering coefficient

$$C_c = \frac{3N_T}{N_3}$$

takes the ratio of the number of triangles N_T in the network with an edge between each node, to the number of connected triples N_3, in which nodes can reach other nodes either directly or indirectly.

Graph types used to simulate P2P networks include the Barabási (BA) preferential-attachment model, capturing the scale-free and small-world properties; the Watts-Strogatz (WS) model, capturing the small-world and clustering properties; and the Erdős-Rényi (ER) random graph model, since in selected undirected P2P networks (e.g. Gnutella) peers choose neighbours essentially at random. We used all three types as no single model captured all P2P properties.

Using the *igraph* package in R, we generated:

1. **Embedded P2P Networks:** 1,000 networks of types BA, ER and WS were generated for the G_C and G_E networks, with the Flight #77 hidden network embedded in each, such that the link connectivity of the hidden network to the remaining nodes was comparable to that in the original G_C and G_E networks. We repeated this process for all remaining hidden networks (Flights #93, #175, #11, Trusted Priors).

Peer-to-Peer characteristics and group-based centralities were calculated for each network, from which averages were produced; and

2. **Simulated P2P Networks:** generated as for the Embedded P2P networks, however without an embedded hidden terrorist network. Averages were taken of the Peer-to-Peer characteristics and centralities.

The degree distributions for the core and extended embedded P2P networks (Figure 4, top row) indicate that the Barabási and Erdős-Rényi networks follow a similar power law. While the actual 9/11 hidden networks also follow a power law distribution, they are visually distinguishable from the embedded Barabási and Watts-Strogatz P2P networks, although more closely mimic the degree distribution of the Erdős-Rényi embedded networks. Comparison between the embedded and simulated networks (Figure 4, top/bottom row respectively) indicates that while there are noticeable differences between the two types of networks, it would be difficult to visually distinguish one from the other, since a power law degree distribution is present in both cases, as would be expected of a genuine P2P network.

The low clustering coefficients (Figure 5) of the embedded (solid line) and simulated (dashed line) Barabási and Erdős-Rényi P2P networks indicate that nodes are connected to only a small number of acquaintances. The 9/11 and Watts-Strogatz networks have larger clustering coefficients, suggesting that the number of nodes directly connected to each hidden network is larger than for the Barabási and Erdős-Rényi networks. While there is little difference between the clustering coefficients for the embedded and simulated Watts-Strogatz and Erdős-Rényi P2P networks, the Barabási networks demonstrate a noticeable difference in coefficient size when the hidden networks are not embedded.

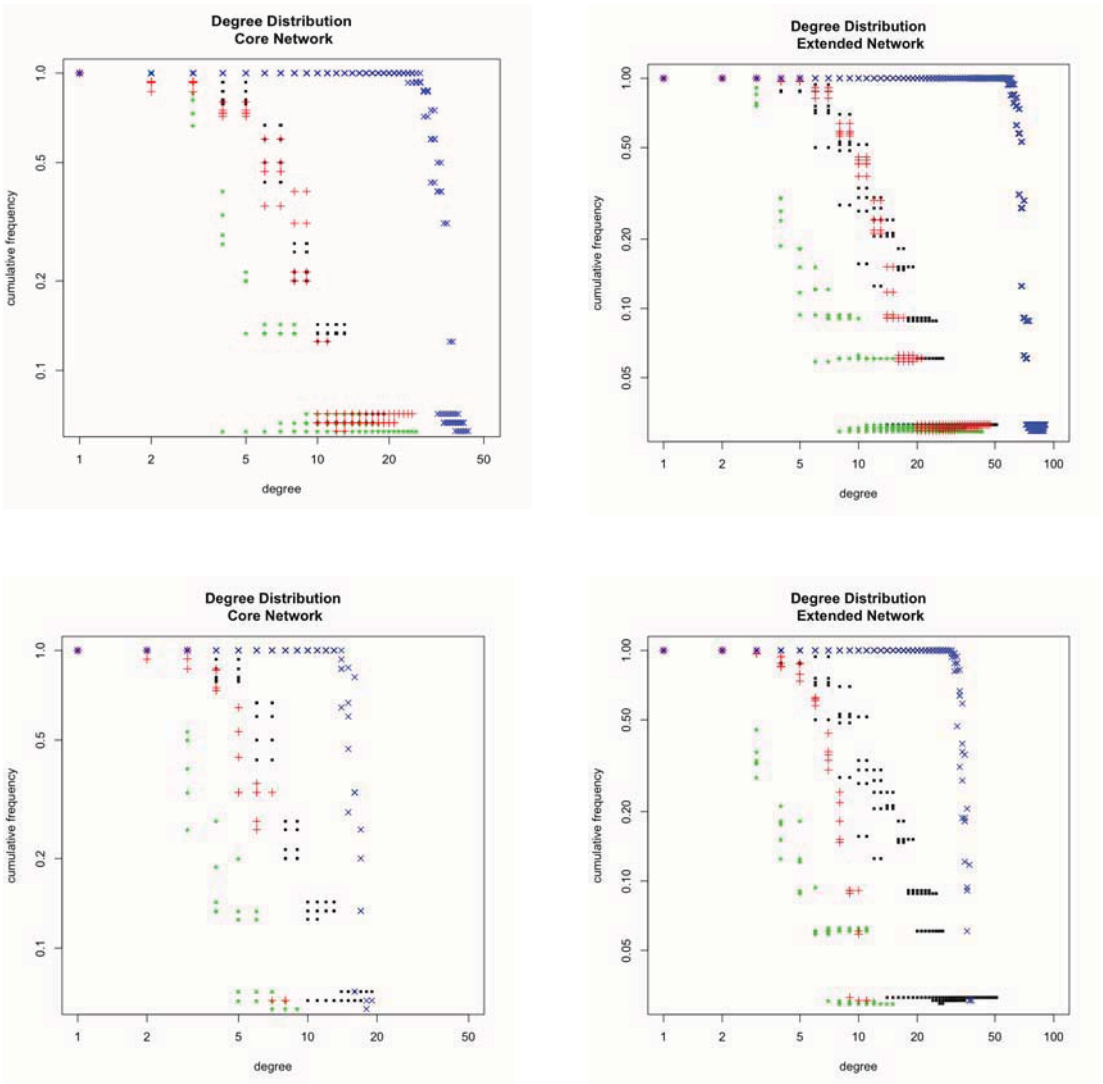

*Figure 4: Degree Distributions. The top row shows all 9/11 hidden networks (black •), and embedded networks using the Barabási (green *), Erdős-Rényi (red +) and Watts-Strogatz (blue X) models. The bottom row shows the simulated P2P networks without embedded terrorist networks.*

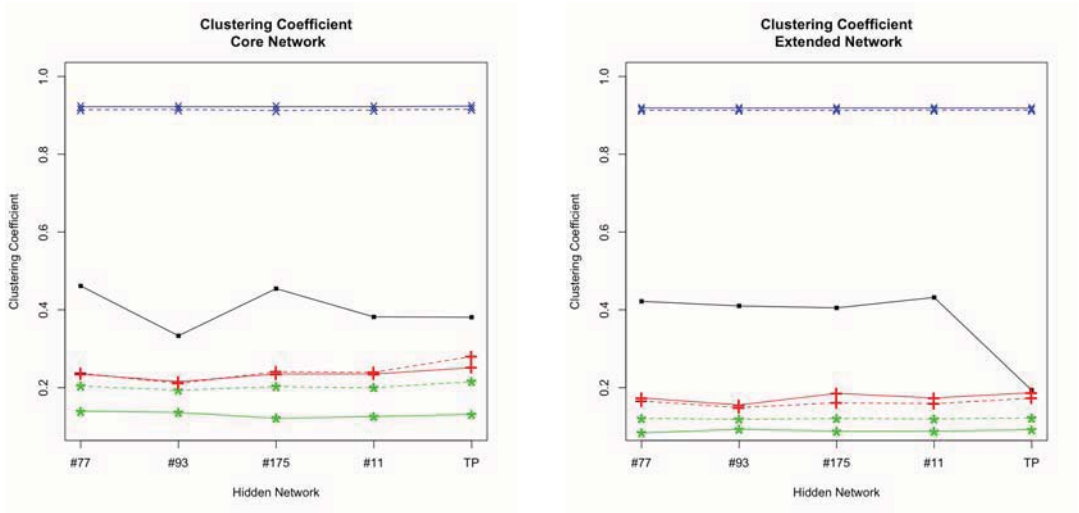

Figure 5: Clustering Coefficients for the Core and Extended networks for all hidden networks of type Krebs (black •), Barabási (green *), Erdős-Rényi (red +) and Watts-Strogatz (blue X). The solid lines indicate the Embedded P2P networks, the dashed lines indicate the simulated P2P networks.

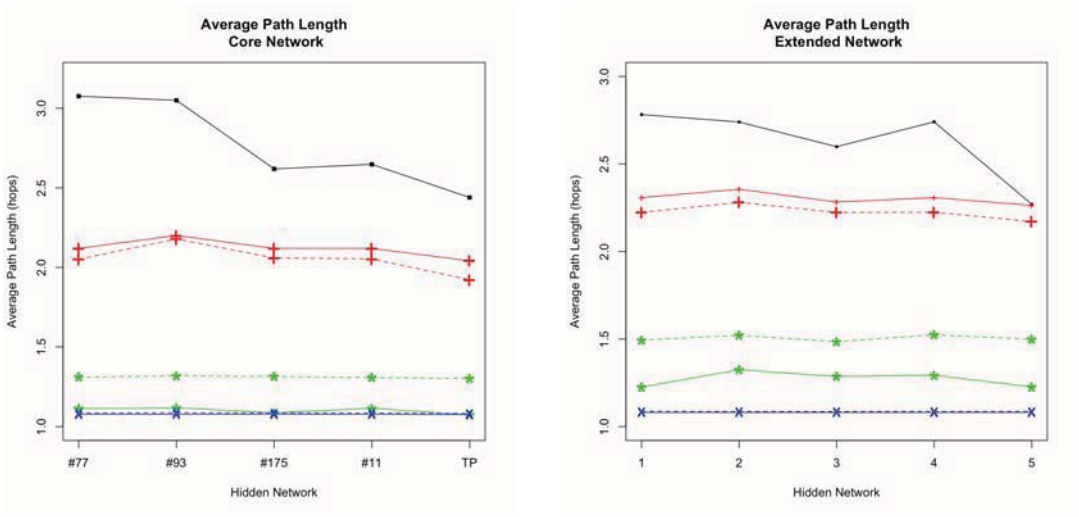

Figure 6: Average Path Lengths for the Core and Extended networks for all hidden networks of type

Krebs (black •), Barabási (green *), Erdős-Rényi (red +) and Watts-Strogatz (blue X). The solid lines indicate the Embedded P2P networks, the dashed lines indicate the Simulated P2P networks.

Similarly, the average path lengths (Figure 6) for the embedded and simulated P2P networks show little difference between the Watts-Strogatz and Erdős-Rényi networks, however the Barabási P2P networks show discernibly different path lengths. The actual 9/11 networks yielded the largest average path lengths, suggesting that although attempting to behave as a P2P network, a surrounding network directly supportive of terrorist activities destroys the average path length property rendering the hidden network visibly distinct from a P2P network background.

The implications of these preliminary results for the Barabási model have untapped potential for locating hidden terrorist networks in cyberspace. For instance it may be possible to detect hidden terrorist networks as non-P2P entities if observed in a P2P network space generally better modelled by a Barabási graph, such as Instant Messaging and Voice over IP (Skype) networks (Newman, 2010).

HIDDEN NETWORK ANALYSIS: BEHAVIOUR

To analyse hidden network behaviour, we computed information, eigenvector and subgraph centralities for the actual 9/11 networks, as well as the embedded and simulated P2P networks (Tables 2-4). To aid interpretation, the 9/11 $\mathcal{G}_C$ network is given in Figure 7.

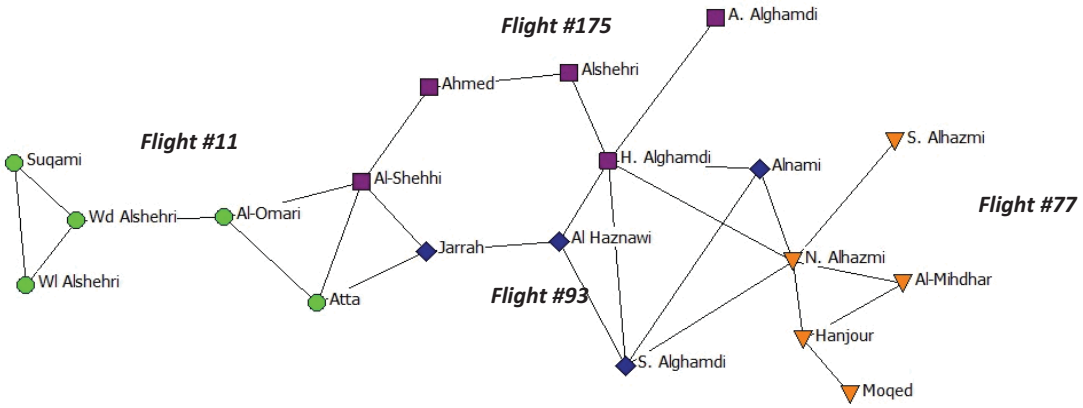

Figure 7: A reproduction of the core September 11 Network of (Krebs, 2002).

Discussion: Information, Eigenvector and Subgraph Centrality Results

Table 2 contains the information centralities calculated for the 9/11 and Embedded and Simulated P2P networks for both network sizes ($\mathcal{G}_C$, $\mathcal{G}_E$). The negative information centralities for the $\mathcal{G}_C$ and $\mathcal{G}_E$ 9/11 networks indicate network inefficiency if the Flight #93 and #175 networks are removed, reflecting the fact that these networks behave as intermediaries in the communication path between the Flight #11 and #77 networks (Figure 7). Thus their removal will detrimentally affect communication across the network.

Table 2: Information Centralities CI for the 9/11, Embedded and Simulated P2P Networks

	#77	#93	#175	#11	Trusted Priors
9/11 Network G_C	***0.15***	***-0.13***	***-0.29***	***0.17***	***0.79***
Embedded P2P G_C Networks					
Erdős-Rényi G_C	0.13	0.08	0.13	0.13	0.23
Barabási G_C	0.82	0.81	0.87	0.83	0.88
Watts-Strogatz G_C	0.12	0.12	0.12	0.12	0.13
Simulated P2P G_C Networks					
Erdős-Rényi G_C	0.17	0.13	0.16	0.13	0.26
Barabási G_C	0.95	0.39	0.25	0.08	0.74
Watts-Strogatz G_C	0.00	-0.10	-0.07	-0.03	0.17
9/11 Network G_E	***0.19***	***-0.12***	***-0.12***	***0.05***	***0.76***
Embedded P2P G_E Networks					
Erdős-Rényi G_E	-0.05	-0.02	-0.08	-0.05	-0.08
Barabási G_E	0.55	0.34	0.50	0.51	0.69
Watts-Strogatz G_E	0.06	0.05	0.06	0.06	0.06
Simulated P2P G_E Networks					
Erdős-Rényi G_E	0.05	-0.02	-0.05	-0.05	-0.08
Barabási G_E	0.85	0.56	0.56	0.50	0.76
Watts-Strogatz G_E	-0.08	-0.20	-0.22	-0.13	-0.21

Removal of the Trusted Priors group however, indicates the overall network will still function efficiently (C_I = 0.79) suggesting the presence of this group is superfluous to communication across the network. None of the embedded or simulated P2P networks produce information centralities in the region of those for the 9/11 networks, particularly the Barabási networks, which understates the importance of each flight network, as indicated by the larger positive values.

The eigenvector centralities (Table 3) capture the importance of each of the flight networks. Recalling that the larger the eigenvector centrality, the more important the hidden network, it can be seen that for both the $\mathcal{G}_C$ and $\mathcal{G}_E$ networks, the embedded Barabási and Watts-Strogatz P2P network centralities predominantly overstate and understate network importance respectively, when compared to the centralities for the actual 9/11 network.

The simulated networks show a better correspondence to the actual 9/11 network centralities. Note however that the results for the embedded and simulated $\mathcal{G}_C$ and $\mathcal{G}_E$ Erdős-Rényi networks indicates a general similarity of importance, to make distinguishing between the embedded and simulated networks a more difficult undertaking. Revisiting the $\mathcal{G}_C$ network in Figure 7, we see that the "better connected" networks, in terms of eigenvector centrality, appear to be the Flights #77, #93, #175 and Trusted Priors networks, since these networks are connected to nodes that are themselves highly connected. The Flight #11 network possesses two unique, albeit inferior, connections, to the remainder of the network, as reflected by the negligible eigenvector centrality (0.06), however when immersed in the larger $\mathcal{G}_E$ network, the eigenvector centrality of the Flight #11 network is vastly increased due to the improved connectivity of group member Atta (Figure 2). It should be noted that Flight #77 is considered more important in the smaller $\mathcal{G}_C$ network (x_i = 0.50) than in the extended $\mathcal{G}_E$ network (x_i = 0.04), due to the inferiority of its connections in the latter instance.

Finally, the subgraph centralities C_s (Table 4) depicts the ease with which information is shared throughout the network. The larger centralities for the G_C Flight #175 and Trusted Priors networks (28.67, 29.56 respectively), indicate that information circulates more easily, seemingly capturing the fact that the Flight #175 network has multiple connections to two of the three other hidden groups in the network. In contrast, the Flight #11 network distributes information poorly, as it is connected to two other hidden networks, albeit through nodes which themselves are poorly connected.

Table 3: Eigenvector Centralities x_i for the 9/11, Embedded and Simulated P2P Networks

	#77	#93	#175	#11	Trusted Priors
9/11 Network G_C	***0.50***	***0.57***	***0.60***	***0.06***	***0.60***
Embedded P2P G_C Networks					
Erdős-Rényi G_C	0.59	0.57	0.59	0.59	0.61
Barabási G_C	0.71	0.71	0.72	0.71	0.71
Watts-Strogatz G_C	0.35	0.33	0.35	0.35	0.37
Simulated P2P G_C Networks					
Erdős-Rényi G_C	0.59	0.54	0.59	0.59	0.61
Barabási G_C	0.87	0.46	0.46	0.42	0.64
Watts-Strogatz G_C	0.39	0.54	0.50	0.44	0.59
9/11 Network G_E	***0.04***	***0.40***	***0.50***	***0.49***	***0.66***
Embedded P2P G_E Networks					
Erdős-Rényi G_E	0.52	0.49	0.54	0.52	0.54
Barabási G_E	0.70	0.69	0.69	0.69	0.69
Watts-Strogatz G_E	0.24	0.22	0.24	0.24	0.25
Simulated P2P G_E Networks					
Erdős-Rényi G_E	0.51	0.45	0.51	0.51	0.54
Barabási G_E	0.71	0.48	0.49	0.44	0.64
Watts-Strogatz G_E	0.24	0.54	0.48	0.24	0.57

The reverse situation is observed for the G_E network in which the Flight #11 and #175 networks are now better located to circulate information more easily, as reflected by the larger subgraph centralities (258.21, 360.99). The inordinately large subgraph centralities of the Watts-Strogatz models for the embedded P2P networks are not uncommon and indicate an abundance of short path connectivity of a network node (Estrada and Rodríguez-Velázquez, 2005). In this case, these large subgraph centralities precisely reflect the observed low average path lengths (Figure 5) and high clustering coefficients for the Watts-Strogatz networks.

Comparison of the subgraph centralities for the G_C Erdős-Rényi embedded and simulated networks (Table 4), indicates once more, the differences between these networks are barely distinguishable from one another. This suggests it would be difficult to detect the presence of a hidden terrorist network in an Erdős-Rényi type P2P network. Conversely, there are somewhat larger differences between the Barabási embedded and simulated networks, particularly for subgraph centrality in the G_C networks. This is encouraging since this suggests that a smaller network background is better suited to detect such a difference, making locating hidden terrorist networks a computationally tractable exercise.

The outcomes of the centrality-based analyses (Tables 2-4) reinforce the observation made earlier in this work, namely that Barabási type P2P networks provide the least protection for a hidden terrorist network wishing to avoid detection. This result paves the potential for future research in hidden network detection in cyberspace, as there are known Barabási type P2P networks that can be modelled for such an exercise. As indicated earlier, Instant Messaging and Voice over IP (Skype) networks provide exciting potential in this context (Newman, 2010).

Finally, it should be noted that the G_C and G_E hidden networks presented here have been characterised in further detail in previous works (Chiera, 2011, Chiera, 2010) in terms of group-based social network analysis including degree, betweenness and induced centrality, as well as network centralisation. The ease with which a hidden network can be detected was determined using these measures, with the analysis suggesting hidden networks are paradoxically easier to detect within larger terrorist-supportive network structures. The interested reader is referred to these works for further details.

Table 4: Subgraph Centralities $C_I(i)$ for the 9/11, Embedded and Simulated P2P Networks

	#77	#93	#175	#11	Trusted Priors
9/11 Network G_C	***11.49***	***11.68***	***28.67***	***6.15***	***29.56***
Embedded P2P G_C Networks					
Erdős-Rényi G_C	28.91	26.02	28.92	28.92	29.99
Barabási G_C	0.05	0.04	0.05	0.05	0.06
Watts-Strogatz G_C	445,009.40	871,280.51	445,009.41	445,009.38	222,822.23
Simulated P2P G_C Networks					
Erdős-Rényi G_C	26.88	26.88	28.22	28.22	30.97
Barabási G_C	23.70	23.70	23.70	23.70	23.70
Watts-Strogatz G_C	15.39	29.97	29.97	29.97	46.57
9/11 Network G_E	***66.51***	***188.37***	***360.99***	***258.21***	***254.74***
Embedded P2P G_E Networks					
Erdős-Rényi G_E	171.13	121.03	181.19	171.14	207.48
Barabási G_E	0.02	0.01	0.01	0.01	0.02
Watts-Strogatz G_E	>1e+13	>1e+13	>1e+13	>1e+13	>1e+12
Simulated P2P G_E Networks					
Erdős-Rényi G_E	228.10	228.10	228.10	228.10	229.08
Barabási G_E	1.73	1.82	1.89	1.90	1.90
Watts-Strogatz G_E	19.03	45.69	45.69	45.69	58.99

Discussion: Assortativity, Limitations and Implications for Future Analysis

Finally, as assortativity provided the motivation for the bulk of the analysis presented here, some discussion on the limitations of assortativity is in order. Whilst it is conceptually straightforward to interpret assortativity based on the sign and size of the value reported, there are known problems with this approach (Newman, 2010). Specifically, perfect scalar disassortativity (r = -1) occurs when the network closely resembles that of a random mixing network, in which dissimilar nodes mix such that the nodes that are *most* dissimilar mix with one another. For degree-based assortativity the conditions for perfect disassortativity become even more stringent. Such a real-world network would be infrequently encountered and it is postulated that most disassortative networks instead produce values closer to $r = 0.0$ than assortative networks (Newman, 2010).

A second point for consideration is whether obtaining a disassortativity value in the region of -0.37 necessarily indicates that a hidden terrorist network behaves as an undirected P2P network. The results presented here suggest some confirmation of this supposition, particularly with regards to the analysis of degree distribution, however a more in-depth analysis into the competing characteristics of P2P networks with hidden networks of varying sizes and configuration is needed before such a statement can be made with firmer conviction. As the networks under consideration expand in size, it would be of interest to determine the sensitivity of the degree-based assortativity coefficient to these changes. Moreover, in the current work it is assumed the connectivity between individuals is undirected and that each hidden network member has equal weighting in the analysis. This may not be realistic, particularly in the case of Skype or Instant Messaging networks, and once such assumptions are removed, assortativity will need to be re-evaluated in terms of these changes.

Finally, it should be noted that in this work we considered a global degree-based assortativity, however there are a number of alternative definitions of assortativity. For example, it is possible to consider *local* assortativity (Piraveenan et al., 2009b) used to capture an individual node's direct contribution to the overall network assortativity. Assortativity can also be defined as a function of time, state or network entropy (Piraveenan et al., 2009a) and it is of interest to determine how alteration of the definition of assortativity will affect the analysis presented here.

CONCLUSION

We investigated whether hidden terrorist networks are more difficult to detect when located within larger non-terrorist networks. Using assortativity, we showed that hidden terrorist networks have the potential to behave as Peer-to-Peer networks, and thus generated embedded and simulated Peer-to-Peer networks for analysis. Quantitative analyses of these networks alongside two actual September 11 terrorist networks indicated that it is possible to distinguish between a hidden network hiding within a larger terrorist network, a Barabási network containing a hidden terrorist network and a genuine Barabási network.

The results of this exploratory analysis hold far-reaching potential for future research in the detection of hidden networks in cyberspace. Future areas of research based on the work presented here include the development of more sophisticated hidden network detection techniques, based on alternative definitions of assortativity. The detection of hidden networks located amongst Skype or Instant Messaging traffic is deserving of attention, as well as the development of models of the evolutionary behaviour of these hidden networks, as compared to that of genuine Peer-to-Peer network traffic, for better detection capabilities. As the goal is to detect these hidden networks in order to stop further terrorist-driven tragedies, it is encouraging that there are still many unexplored opportunities in this area.

REFERENCES

Baumes, J., Goldberg, M., Hayvanovych, M., Magdon-Ismail, M., Wallace, W. and Zaki, M. (2008) "*Discovering Hidden Groups in Communication Networks*". In Gal, C.S., Kantor, P.B. and Shapira, B., *Security Informatics and Terrorism: Patrolling the Web – Social and Technical Problems of Detecting and Controlling Terrorists' use of the World Wide Web,* 15, 82-108.

Borgatti, S.P. (2003) "The Key Player Problem". In: *Breiger, R., Carley, K., Pattison, P.(Eds.), Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*. National Academy of Sciences Press, 241-252

Chiera, B.A. (2011). "Towards the Localisation and Detection of Hidden Terrorist Networks : A Group-Based Social Networks Analysis". *In: Proc. of the 2010 National Security Technology and Research Conference,* Canberra, ACT, 22-23 September 2010, 5-20.

Chiera, B.A. (2010). "Group-Based Social Network Characterisation of Hidden Terrorist Networks". *In: Proc. of the 1st Intnl Cyber Resilience Conference*, Edith Cowan University, Perth, Western Australia, 23rd August 2010, 11-21.

Estrada, E. and Rodríguez-Velázquez, J. A. (2005). "Subgraph centrality in complex networks". *Phys. Rev. E, 71(5)*, 0561031 – 0561039.

Farall, L. (2011). "How al Qaeda Works: What the Organization's Subsidiaries Say About Its Strength". *Foreign Affaris,* 90, 128-139.

Freeman, L.C. (1979). "Centrality in social networks: Conceptual clarification". *Social Networks* 1, 215-239.

Hussain, D. M. A. (2010). "Investigation of Key-Player Problem in Terrorist Networks Using Bayes Conditional Probability". *In: Furht, B. (Ed.), Handbook of Social Network Technologies and Applications*. Springer New York.

Krebs, V. (2002). "Mapping networks of terrorist cells". *Connections* 24, 43-52.

Lindelauf, R., Borm, P. and Hamers, H. (2009). "The influence of secrecy on the communication structure of covert networks". *Social Networks* 31, 126-137.

Newman, M.E.J. (2010). "Networks. An Introduction." Oxford University Press, New York, NY.

Piraveenan, M., Prokopenko, M. and Zomaya, A.Y. (2009a). "Assortativeness and information in scale-free networks". *Eur. Phys. Jnl B,* 67, 291-300.

Piraveenan, M., Prokopenko, M. and Zomaya, A.Y. (2009b). "Local assortativity and growth of Internet". *Eur. Phys. Jnl B,* 70, 275-285.

Wang, F., Moreno, Y., and Sun, Y. (2006). "The Structure of Peer-to-Peer Social Networks". *Phys. Rev. E* 73(3), 1-7.

Wasserman, S. and Faust, K. (1994). "Social Network Analysis: Methods and Applications". Cambridge University Press, 857 pages.

Watts, D.J. and Strogatz, S.H. (1998). "Collective dynamics of 'small world' networks". *Nature* 393, 440-442.

Wu, Z., Jiang, G., Zhang, C. and Tang, Y-Y. (2010). "Traffic Organization Method for Emergency Evacuation Based on Information Centrality". *Second International Conference on Advanced Computer Control,* Shenyang, China, 27-29 March, 92-96.