

RADICALIZATION ONLINE

Patterns of Social Interaction on the Al-Faloja and As-Ansar Forums

Lucas Lenselink
3060438
Utrecht University
12 August 2011

A Thesis submitted to the Board of Examiners in partial fulfilment
of the requirements of the degree of Master of Arts in Conflict Studies & Human Rights

هماسا دي اقل

Name of Supervisor

Chris van der Borgh

Date of Submission

12 August 2011

Programme Trajectory

Research & Thesis Writing only (30 ECTS)

Word Count

21,184

Opgedragen aan mijn moeder, die haar thesis aan
mij opdroeg toen ik nog geen jaar oud was.

LL

Thank you,

The Contributors to the Dark Web Forum Portal,
Catherine Larson with the Artificial Intelligence Lab, University of
Arizona,
Sascha Ramondt,
Chris van der Borgh,
CCS Staff.

Cover design: Myrthe Lenselink

Contents

- 1. Introduction: a Story of Radicalization through the Internet**
- 2. Current ideas in Social Sciences on Terrorism, Internet, and Radicalization**
- 3. Quantitative Methods for Research on the Internet**
- 4. Social Network Methods for Studying Large Sets of Data**
- 5. Methodology**
- 6. Findings**
- 7. Conclusion**

Appendices

Bibliography

1. Introduction: a Story of Radicalization through the Internet

Terrorism, it seems, is a timeless concept. It is not limited to geographical borders or to a religious denomination. It is a concept that keeps perpetuating itself. In war-torn regions in the Islamic world as well as in modern western welfare societies individuals with radical ideas have developed themselves to becoming terrorists. A very recent example is Anders Breivik, whose actions in Oslo and on the island Utøya were watched throughout the world with disbelief and disgust. London, Madrid, Amsterdam, New York are all cities that have been confronted with the sick deeds of a brainwashed individual. Yet all perpetrators or terrorists once were healthy and stable human beings. How has it come so far? A desire to try to understand this path of radicalization underlies this thesis. As an example of such a path I would like to introduce the story of Mohammed Bouyeri, a Dutch-Moroccan Islamist, who is currently serving a life-sentence for assassinating the Dutch film director Theo van Gogh.

On the 2nd November 2004 cineast and criticaster Theo van Gogh was murdered by Mohammed B., a Netherlands born and raised man of Moroccan descent. His motivations were religious and political. This brutal act of terrorism shocked people all over the Netherlands. Everyone was talking about it, had an opinion about it. There was no way to withhold from the discussion. Reactions were both of muscled rejection and emotional disgust. How could someone be so frustrated as to kill a rude and sometimes even vulgar, but innocent person? Mr van Gogh (born 1957) was widely known as an enfant terrible, a provocateur, as much as he was known as a filmmaker. His rude and vulgar criticisms on the government, the multicultural society and Islam gave him a reputation of an unbettable warmonger. He had many friends, but even more enemies.

Mohammed B. grew up in Amsterdam-West. He went to primary school and played football on the square. He was a shy person, especially with girls. Other than many of his peers he performs as good as he can in school. Both his teachers and his fellow pupils have a positive image of him: a timid but clever boy that wanted to make a career. In his neighbourhood there is a lot of nuisance from youth, but Mohammed B. behaves exemplary. He is active in the youth centre '*De Oostoever*'. Considered a natural born leader he succeeds in getting some of the youth off the streets. However in 1994 the youth centre is demolished and the migrant youth return back to the streets. Mohammed B. feels betrayed and not taken seriously. Up until then, Mohammed is not a practising Muslim. He loves to drink beer and uses soft drugs. He decides he wants to become an accountant and enrolls in the

Hogeschool INHolland, together with a friend. Studying is not a success. He switches several times between different studies, and eventually leaves the school after five years without any diploma. Meanwhile he also begins to show fanatic traits, for example harassing young Muslims who drank alcohol¹. Mohammed continues to work on the problems in his neighbourhood. He keeps advocating for a youth centre. The local bureaucracy processes his initiatives very slowly. His blocked ambitions fed his frustration, and he comes into contact with the police several times. Mohammed gets arrested for involvement in bar fights and extremely aggressive behaviour toward police officers and is sentenced to 12 weeks in prison. While in prison, the Islamic faith is starting to become important to him and he starts studying the Quran. When released from prison in September 2001, Mohammed encounters even more problems. His mother dies of breast cancer and his father has severe back pains and must cease working. His numerous propositions for improving the position of migrant youth in the neighbourhood are repeatedly rejected. Mohammed is deeply disappointed: years of hard work undone with one fell swoop. He quits his education, without a diploma, and becomes dependent of social welfare provision.

It is then when he turns to radical Islam. He starts looking for radical Islamic texts on the Internet that appeal to him. He is completely absorbed by this quest for meaning. His appearance is also changing: he grows a beard, wears a traditional djellaba, loses his openness and joviality and grows cold and distant instead. Despite getting isolated more and more, the neighbourhood centre offers him a job as janitor. His employment is very problematic as he insists on Islamic norms and values being abided in the centre – he refuses to accept that people drink alcohol within the centre, insists on the separation of men and women. Efforts to talk with him are pointless as he refuses to alter his ideas. His employment therefore is ended, and Mohammed becomes even more isolated, spending entire days behind his computer reading, writing, translating and distributing radical Islamic texts via the Internet. He actively starts participating as an online operating Muslim extremist.

Mohammed also hosts meetings at home with radical re-Islamized youths, a network in the Netherlands commonly known as the ‘Hofstadgroep’. Radical imams and other Islamic clerics are invited to give lectures. They are prepared for the jihad, and Mohammed is thus alienated more and more from his (former) friends and family, but also from his clergy. He visits more radical and salafi²-oriented Islam communities throughout the country. Eventually even the Al Tawheed mosque, which is believed to be among the

¹ Trouw, 9-7-2005.

² ‘Salafi’ means ‘predecessors’, which refers to the earliest generations of followers of the prophet Mohammed. Its adherents are known to be very puritan and orthodox.

most orthodox mosques in the Netherlands, is too liberal for Mohammed. Being isolated from his peers, and interacting mainly online with equally minded people, his radical views are encouraged even more by his direct environment and so he is very much convicted of his justness. He even visits imams and announces that he ‘will show them what Islam is’³. By December 2002 Mohammed already proposed bombings on a large scale with many deaths, according to one of the leaders of the ‘Hofstadgroep’. He embraced the call for a violent jihad, and spread this call actively on diverse forums on the Internet and via e-mail.

On the 1st November 2004 Mohammed dines for the last time with his friends, knowing that it would be the last evening of his life. The next morning Mohammed leaves the house early. Theo van Gogh is shot at with at least 20 bullets, after that he is stabbed aggressively and full of hate. Finally Mohammed cuts his throat and leaves his testament with a kitchen knife on the breast of his victim.

What ever happened with the gentle, intelligent and servile Mohammed? What experiences drove him in the direction of radical and violent Islam? The story of Mohammed B. calls up questions about the influence of online social media. Such media might have played an influential part in the radicalization process of Mohammed and his friends, as it is known that they intensively made use of Internet services, e.g. MSN Groups, to shape and propagate their ideas (Benschop 2005). Another example of such sites is the website hosting company 357hosting.com. This company, established in Nieuwegein (The Netherlands), was in 2004 considered to be the world’s largest terrorist-host, supposedly supporting thousands of terrorist web sites. 357hosting.com specialises in the anonymous hosting of extremist and terrorist web sites (ibid.). The company gives large discounts to websites with Islamic content, giving rise to the suspicion that wealthy fundamentalists support it. Most of these websites have been shut down by (Dutch) authorities.

Investigation by Dutch secret service did not estimate Mohammed to have the greatest terrorist potential. Some of his friends (or associates), and their Internet activity, however, have been closely examined. Bilal L., for example was active on numerous MSN groups (Al-Ansar, Shareeah, A Salafoe Saali7 en 9113 (Benschop 2005)) and web sites such as marokko.nl and maroc.nl, where he among other things, posted addresses of flying schools, shooting clubs, and advices on Islamic literature. His comments, texts and advices are of a very practical nature, and have a highly mobilizing character. Consider his advice on how to behave when practising on a shooting range:

³ NRC Handelsblad, 12-11-2004.

‘Try to keep this kind of activities a secret, so that you may depart from a pure *niyya* [intention; my italics]. At the shooting range, keep your opinions and convictions to yourself, and don’t go in conclave with the others that are present. Don’t talk about Islam and perform your *salaat* [Islamic prayer; my italics] in secret. You are going to the shooting range as a preparation for the Jihad and not to invite people to Islam.’⁴

There are shooting clubs in the Netherlands that allow their members to take their firearms home after a certain period of membership. Bilal moves on to advice:

‘Do not do this if you cannot control your aggression, or if you have problems in your personal life. Respect the Dutch legislation and avoid buying illegal firearms. There are ample opportunities to train legally, so don’t ruin your reputation by turning the illegal way. Learn most what you can learn in your own society and learn the rest when you actually arrive in a country of the Jihad.’⁵

He rudely criticized more moderate Muslims, and approves of (extreme) violence against Western people. More specifically he calls for the assassination of Theo van Gogh:

‘It is an obligation to kill he who abuses the Prophet whether he is Muslim or Kaafir. And Hirsi Ali [Dutch liberal politician and former Muslim] and Theo van Gogh, these pigs who have abused the prophet their punishment is death and their day will come with Allah’s will..!’⁶

He reinforced his call with some texts from the 14th century that were translated by Mohammed. Bilal is thus paving the moral way for the actual murder. Bilal also called for the Dutch politician Geert Wilders to be beheaded. He was arrested and convicted, because web sites like the MSN group where he posted this calls get a lot of public attention, and so his words had a terrorist intent. Bilal declared that he got interested in the jihad through the Internet. Imams in Amsterdam could not answer his questions and so radical forums were the only place where he could get answers.

Bilal also declared that he got caught up in the unrestrained rhetoric of Internet forums. Indeed, there is little, if any, social control on those forums. It is very easy to call all kinds of extremist-jihadist creeds, when the social establishment favours the extreme, as is the case within such forums. The question remains, to what extent do such statements lead to real action? What is the mobilizing value of for example Bilal’s posts? And do ‘keyboard-terrorists’ become real-life terrorists or do they inspire others to do so? Both statements from convicted (internet-)jihadists and from Dutch secret service show that the mobilizing effect of Internet forum posts should not be underestimated. Bilal of course was not the only one of Mohammed B.’s

⁴ Posted under pseudonym Abu Qataadah, original post deleted from site, but see repost at <http://forums.marokko.nl/archive/index.php/t-206337.html>, post translated by Benschop.

⁵ Ibid.

⁶ Ibid.

friends that was active on the Internet propagating the jihad in the Netherlands. Another acquaintance of his was Omar A. After the release of the anti-Islam film ‘Submission’ by Ayaan Hirsi Ali and Theo van Gogh many forums and online groups boasted strongly violent posts. Although shocking, it was nothing new. What was new, however, was that Omar under the pseudonym of Abu Nawwaar revealed the home address of Hirsi Ali. This heavily alarmed Dutch secret service: the address was correct! This shows how online discourse had gone beyond words; also it points at the dangerous potential of online forums.

At that time, however, this potential was still not fully recognized by Dutch secret service. They were aware that potential terrorist intensively made use of the Internet, but not of the coordinating, persuasive, and radicalising role the Internet played. It is the latter role of the Internet that I am interested in. Among other reasons, this was due to a lack of technical data collection tools. Relying mostly on traditional, locally gained data, they failed to effectively connect their ‘traditional’ data with more technically advanced data collection techniques for surveying the Internet (Benschop 2005). The lack of knowledge is put well by American scholar Robert Pape, who argues that the fact ‘[t]hat Internet has become increasingly important in the dissemination of radical Islamic thoughts and sentiments has become a cliché by now. *How* they use Internet is a lot less clear. Islamic fundamentalists and terrorists do not operate in formally and rigidly organized groups, but rather behave like swarms of bees, capable of stinging from all directions. (...) [D]ue to its low costs of access, its speed and worldwide scope, Internet has become the largest megaphone in the world. Besides, it is a public space where you can say what you want, without anybody noticing immediately who you really are. After all it doesn’t take long before one meets a number of people of the same mind. (2005, cited in Benschop, my italics.)

Nonetheless, Mohammed by that time had published numerous Islamic texts that he had written or translated. Additionally, he posts column-like pieces in which he analyses the situation in the Netherlands. His language is fairly good, his style forceful, one-dimensional and dogmatic, and his ideas violent and cruel. He repeatedly puts emphasis on the actuality of the jihad and the call for all ‘upright Muslims’ to stand up and join it:

‘Free yourself! Come out of that coffee shop, come out of that bar, come out of that corner. Answer the appeal of *la ilaha illa allah* [‘There is no god but Allah’]. Join the caravan of the Martyrs. Rise from your deep sleep, rise and shake off the dust of humiliation. Rise and answer *hajja al jihaad*’s call⁷

⁷ From Mohammed B.’s own text ‘To catch a wolf’, see <http://www.freewebs.com/overigeinfo/eenwolfuitschakelen.htm> published 16-04-2004.

The essay is closed with a defiant note: ‘Mr Donner, what will be your next bill to stop us terrorists?’⁸ Still the issue is not taken serious by Dutch secret service. Mohammed moves on to write threatening letters to Dutch politicians with a notable position towards Islam, for example because they are former Muslims or have a strong anti-Islam position. At some point, Mohammed must have decided that it was he himself that had to perform this ‘sacred’ task. From the preparations he made prior to the attack we know that he was clear of mind and that he worked very focused. His exact motives, however, remain unknown.

This horrific story shows the impact the Internet can have on a heavily disappointed and frustrated young man, driving him to ritualistic slaughter. The story of Internet-facilitated radicalization through Islamic forums, however, does not end here. You could imagine that radical sites affiliated with the Hofstadgroep would have been removed from the Internet. On the contrary: within a week after the murder a hitherto unknown pro-Al Qaeda groups released new threats for bombings in the Netherlands in retaliation of the arson of an Islamic school in Eindhoven. Also, a new MSN-group called ‘*De Basis*’ (a literal translation of the name of ‘Al-Qaeda’) appears on the Internet. The web site contains documents, videos, pictures, letters and stories in which the jihad openly is embraced and encouraged. ‘De Basis’ is not the only one. Benschop (2005) lists 22 more Dutch jihad web sites that were active around that time. Those websites just pop up and disappear, as it is very easy and free to create them. The above story is just an example of a worldwide trend: a sharp rise in both the number of active terrorist websites and the richness and sophistication of these websites. Worldwide it is estimated that by the end of 2006 there were 5,300 terrorist websites (Weimann 2008: 75). Increasingly the content of those websites becomes more and more sophisticated with more multimedia elements.

A few factors from the story of Mohammed stand out. A promising student became a frustrated citizen; but the trajectory from frustrated citizen to terrorist is what this thesis is about. More specifically: the part of that trajectory that happened on the Internet. Of course, being an extremely disappointed and eventually frustrated citizen who feels let down by society time after time made Mohammed vulnerable and susceptible to radical Islam. But the influence of social interaction with peers on radical Internet forums, of which this story provides some anecdotal evidence, appears to have been enormous. Radicalization seems to have taken place group-wise. Mohammed and his friends, it appears, together moved along on the path of radicalization. The Hofstadgroep seems to have been a close-knit group. Also, it seems that the process of online radicalization is one of intimidation and influencing of powerful radical actors. This has sparked my interest: the social structure of Internet forums. How is it that radicalization takes place

⁸ Ibid. Mr Donner at the time was minister of the Interior.

online, also in other instances than this Dutch example? My assumption is that the dominant structure of radicalization is clique-wise, fed by external influential actors.

In order to more closely research and verify this assumption, I have been able to monitor the world's most important extremist and terrorist forums through the Dark Web Forum Portal, created by experts from the University of Arizona. Working with advanced software the entire World Wide Web can be monitored for terrorist activity. I have been granted graded access to the Portal, which enables me to search for specific radicalization-related keywords. Using methods and techniques from social network analysis I have been able to process this gigantic amount of data and distill some very useful information on social structures and patterns of social roles. This thesis is a detailed account on how I have done this, on the underlying ontology, and on the subject population.

I will first elaborate on current scientific insights on the subject under scope: extremists and terrorist organizations. I will zoom in further on one specific branch of terrorism, the Global Salafi Jihad and their presence on the Internet. They use the Internet for a couple of distinct purposes, one of them being 'spreading the word': propagate their ideology and seeking adherents. This is not one-way traffic but rather a process of interaction and that is why online forums and chat rooms are extremely important: geographical boundaries become irrelevant, and the direct online social environment approves of extremism and even encourages it, as online communication is direct and unrestrained. Using more advanced techniques from the field of Informatics a number of forums have been identified as being hotbeds of radicalization. From the content of these forums we can learn a lot about the social structure, that shapes the climate for radicalization. Gaining a better understanding of that social structure of radical Islamic forums is the main goal of this thesis. I will discuss the methods that have been used identifying and collecting these forums, as well as the methods that I will use to abstract social structures. Finally I will present my findings and interpret them.

2. Current Ideas in Social Sciences on Terrorism, Internet and Radicalization

In this chapter, I will try to give an outline of what current ideas in social sciences are on Internet, terrorism, and radicalization. These days Internet is used by almost everyone in our environment. Internet, as some argue, is ‘in many ways an almost perfect embodiment of the democratic ideals of free speech and open communication; it is a marketplace of ideas unlike any that has existed before’ (Weimann 2008: 82). And if we look at some key characteristics of the internet – its decentralized nature, the fact that it cannot be subjected to control, restriction, or whatever censorship, and its easy accessibility to everyone – I tend to agree with them. However, it is not too hard to imagine that it is these very characteristics that make the Internet also vulnerable to abuse by less benevolent users. What’s more, today all known active terrorist groups have established at least one form of presence on the Internet! (Weimann 2004: 1) The study of Internet use by this kind of groups is very young. Also, the subject field is very dynamic, as websites appear and vanish sometimes in as much as a few days. Hence the number of relevant literature on terrorism and Internet is limited.

I will first explain what I think ‘terrorism’ and ‘extremism’ mean, so as to demarcate the phenomenon that I am theorizing about. I will then zoom in on a particular branch of terrorism, the Global Salafi Jihad. I will then move on to explain what ‘radicalization’ is, how it can be studied, and what it may lead to. Then I will elaborate on the presence of the groups involved on the Internet, and shine some light on the question why the Internet is so appealing to them. Finally I will discuss current analysis and theorizing in this relatively young field, and explain why conventional data collection techniques are not adequate enough.

Terrorism

This thesis is about the use of Internet by terrorists. By ‘terrorists’ I mean persons, members of groups associated with the use of terrorist strategies and tactics. That is, the use of violence against other persons who are targeted indiscriminately rather than because they are a member of an opposite group. The purpose of the attack is not to eliminate an enemy but to instil fear in a large group. Stohl defines terrorism as ‘the purposeful act or the threat of the act to create fear and/or compliant behaviour in a victim and/or the audience of the act or threat’ (2007: 229). The word ‘terrorist’, stemming from the Latin *terrere*, to frighten, is a very normative term and is always used in a condemning sense. No one will refer to himself as being a terrorist. ‘Terrorist’ refers, however, to the strategies and tactics employed and not to the type or composition of the group. So whenever in this review I use the word terrorist, I mean these kinds of groups rather than

condemning what they are doing, which for the purpose of this paper would not be interesting.

Some authors, especially within the security informatics domain, also write about extremists. This is not exactly the same as terrorists, although there are a lot of similarities. I consider terrorism to be an act, while extremism is a belief. Internet is an ideal means for spreading a message. Spreading ideology is the most important aim of terrorist use of Internet (Chen et al. 2008: 1357). The same applies for extremists. An additional important aim of terrorist Internet use is also recruitment, and there lies the difference: terrorists and extremists both seek to spread a set of beliefs that most people would consider to be radical, but terrorists do so by eventually mobilizing people into violent acts. Terrorism refers to the end goal of a process, not to the current status.

Certainly, most terrorist can also be considered extremists, but not vice versa. A number of authors specialised in web mining research write about extremists instead of terrorist, for example because they aim to address a larger group. Neo-Nazis, for example, are extremists but not necessarily terrorists. When wanting to theorize about the whole field of the spread of radical messages on the net, the distinction between terrorists and extremists becomes less relevant. The distinction *for this purpose* is not so important. That is because I will focus on the process of radicalization. It is about spreading a radical message, and social structures that facilitate such messages to be spread. Let me be perfectly clear: what happens on radical Islamic forums is not terrorism. It is *not yet* terrorism, at most. Radicalization is not terrorism, since it is not directed at a large indiscriminate group. Radicalization can, however, be fuelled by organizations with a terrorist intent. When I speak of a terrorist forum, it will most likely mean that the forum is administered by such an organization. But the content of the forum is extremist, not terrorist. I will use both terms, so please be aware of the difference.

Global Salafi Jihad

A very well-known branch of terrorism is the global salafi jihad. It is radical Islamic. ‘Salafi’ means ‘predecessors’, which refers to the earliest generations of followers of the prophet Mohammed. Its adherents are known to be very puritan, and accept only the Quran and statements by Mohammed himself and his followers as authoritative and is therefore one of the most orthodox currents within Islam. It is in itself not violent, but it is generally intolerant toward other conceptions. A small part of its followers, however, turn it into a violent ideology, by combining salafism with ‘jihad’. The word ‘jihad’ means struggle and has two meanings. First, the inner, or greater, jihad is an inner struggle of the Muslim against the everyday

temptations of life and for submission to Islam. The outer, or lesser, ‘jihad’ is a violent struggle against the enemies of Islam.

The global salafi jihad is far from an organised enterprise. There is little unity within, if present at all. Rather, global salafi jihad is a social movement consisting of a set of more or less formal organizations. Although decentralized, the global salafi jihad is one of the most fanatic and determined ‘branches’ of terrorism. Its actions are mainly targeted at the West, but they indiscriminately hit other races and religions as well. Al-Qaeda can be seen as its ‘vanguard’ (Qin. et al., 2005: 292). The different loosely bonded organisations are linked in patterns of interaction ranging from the fairly centralized to the more decentralized and with various degrees of cooperation, resulting in more or less connected terrorist operations (Sageman 2004: 137). The global salafi jihad is a so-called small world network, strongly clustered groups with a tendency to be exclusivist: having many connections within, but few outside the group. Because of its dense interconnectivity it resists fragmentation well. The structure of the network will not be affected by random attacks. The revolution in communication technology in the 1990s that continues up to date coincided with the rise of the global salafi jihad. Advances in new media and communication technology increase the organisation’s potential as well as its vulnerability, especially to precisely targeted attacks on its hubs, that is, on its leading figures.

Radicalization

So what is radicalization, be it online or offline? According to Sageman (2008) it consists of an interplay of four elements that reinforce each other and together lead toward radical ideas. Sageman’s *‘Leaderless Jihad’* is probably the most well known text to explicitly address online radicalization (Bermingham et al., 2008: 1-2), so I highly rely on that one. Radicalization starts with what he calls *‘moral outrage’* (Sageman 2008: 70). Applied to the case of Mohammed B., an example of this could be his severe disappointment in the Amsterdam society, which repeatedly rejected his plans for better facilities for migrant youths. This moral outrage is then *interpreted in a specific way* (ibid.). With Mohammed, this disappointment or outrage was channelled in radical Islam, especially on extremist online forums, where he did get the affirmation he was looking for. The third step is that the interaction with others, in this case like-minded radical Muslim youth on forums, *resonates with own personal experience*, giving further justification and self-confidence (ibid.). Finally, people may be *mobilized through networks* (ibid.). Although this suggestion of a ‘path to radicalization’ may not always be applicable, some aspects of it definitely make sense. The interaction between disappointment and insecurity on the one hand, and re-interpretation of personal experiences in a secure, but radical environment on the other, is backed by many practical evidence. Indeed, ‘in many of their

writings and speeches, terrorists refer to the strong motivational effects of hearing about or watching the suffering of fellow Muslims (...) brought about by human hands and seen as a major moral violation' (Sageman 2008: 72-73). Another definition, specifically aimed at online radicalization calls it 'a process whereby individuals, through their online interactions and exposure to various types of Internet content, come to view violence as a legitimate method of solving social and political conflicts' (Bermingham et al., 2008: 1).

The process of radicalization can be studied in multiple ways. There are those who follow a biographical approach, focusing on social psychological processes, significant traumatic events in a person's life et cetera, which have driven persons towards terrorism. This micro-level analysis is currently the most common approach. The advantage of the micro-level of analysis is that it is very detailed and gives a thorough insight into the motives of the individual. It is a bottom-up approach. It provides anecdotal evidence, which easily appeals to people, because you can empathize. However, as Sageman rightly points out, a number of problems occur with the micro-level (2008: 16-18).

First, it is difficult to tell how significant, urgent or widespread terrorism is from merely looking at individual terrorists. Second, micro-level analysis underlies the assumption that terrorists are fundamentally different from other human beings, which would mean that there would be such a thing as a terrorist personality or even a terrorist-gen (Sageman 2008: 21). However, even after many decades of psychological research, no such thing has been found. Related to this, it is also assumed in the micro perspective that people are fully aware of what they are doing. Interviewing people is popular as a method to discover the motivations and intentions of people. However, there is overwhelming evidence – from famous psychological experiments as Milgram's and the Stanford Prison experiment, as well as from experiences with perpetrators of human rights violations – that this data collection technique can be highly misleading. People are conditioned to participate in atrocities along a continuum, often without being aware of it. Even if a terrorist is entirely honest with the researcher in an interview, he is very likely not to know what exactly drove him to. Third, micro-analysis focusing on the individual tends to neglect situational factors. Finally, as Sageman argues, 'focusing exclusively on terrorists can never tell us how they are different from those who might have become terrorists but did not' (2008: 22). Micro-level analysis lacks a frame of reference so as to make a comparison with a significant control group.

Others look for root causes of terrorism in the social and social-economic conditions of a society. This is a more top-down look. The macro-perspective is especially prominent among policymakers, because the social, economic and political conditions of a society are what they can affect,

rather than personal choices and preferences (Sageman 2008: 24). The problem with this perspective is that these same conditions are acting on millions of people, while only a minuscule proportion gets involved in terrorism. The macro-perspective does not answer the question why that is. It fails to explain how these conditions are then perceived and interpreted in such a way that it drives some towards terrorism. Also, when conditions happen to be ‘right’, still an overwhelming majority will not get involved in terrorism. How is that? Terrorism seems not to be the inevitable outcome of societal forces, nor the inevitable result of the beliefs and perceptions held by terrorists.

A middle ground, a meso-perspective would therefore be a good alternative. A meso-perspective focuses on how people in groups influence each other to become terrorists. Radicalization has been shown to usually occur group wise. People usually join the jihad in small groups of friends, who can vouch for each other. These kinds of groups are called *cliques*. In abstract terms, in a clique every person is connected to every other one. Cliques play a crucial role in transforming potential contributors into full-fledged mujahedin (Sageman 2008: 154). Indeed, ‘in-group love rather than out-group hate seems a better explanation for the behaviour of terrorists’ (ibid.). Cliques are a way of overcoming the free rider problem⁹. And so it would make sense to study radicalization from this point of view.

So, what does such a group look like and how is it formed? In many terrorist or extremist groups, there is some kind of ‘mastermind’, someone who is older than the rest and speaks Arabic. Those people are suspected of having a strong influence on these groups of young people with forceful language but also with narratives of heroism, as well as mutual reinforcements. A typical example of this is found on the Shamikh forum, on thread #16505 ‘Project to recruit supporters of the jihad (...)’, where a call is made to join the violent jihad¹⁰. Both religious texts and practical tips are found here. Positive replies are praised in all kinds of ways giving a strong impetus on actively or passively supporting the jihad. This virtual global jihad is part of the so-called third wave of terrorism, consisting mostly of ‘homegrown’ terrorists, who start as ‘wannabes’, often without any radical religious background (Sageman 2008: 79). Radicalization is about adhering new ideas through processes of influencing, intimidation, and manipulation: it is about discourse and interaction.

⁹ The free rider problem, in short, is a concept in social sciences which refers to the general social tendency of people to not participate in an enterprise with high risks (such as terrorism), even though they may agree with the ideals of it, but rather let others make the effort and benefit from their endeavour.

¹⁰ DWFP, Shamikh thread #16505

To what may radicalization than eventually lead? For this we need to further zoom in on a particular group, on personal level. Qin et al. in 2005 performed a network analysis of the persons in the global salafi jihad network, in terms of topological properties and organizing principle. A social network analysis, as it is called, is meant to analyse a network in the form of a graph. From such a graph one can distil the ‘heavyweights’, or the ‘gatekeepers’, by measuring the centrality of persons in the social graph¹¹. The Global Salafi Jihad network at that time consisted of 366 terrorists. The outcome of the analysis suggested that ‘the GSJ network may function as a “holding company” model, with Al-Qaeda as the “umbrella organization” (Qin et al 2005: 301). Such a model makes efficient planning possible and reduces the risk of being disrupted, as the operations are carried out by smaller groups that have minimum interaction with the central members of Al-Qaeda. The communication is run via Internet; the internal as well as the external communication.

Internet

Indeed, an important medium through which radicalization is very much happening these days, is the Internet. Generally we can speak of two different ways in which terrorists can make use of the Internet. First, actively or offensively, think of cyberwarfare in terms of hacking and the like. It has been defined by Weimann as ‘the use of computer network tools to harm or shut down critical national infrastructures’ (2005, cited in Stohl 2007: 229) e.g. energy, transportation or government operation. This type of Internet use is widely reported upon in the media. Both fear for technology and fear for terrorism instill a fascination for practically any kind of computer attack among journalists and the public (Stohl 2007: 225). The media attention notwithstanding, cyber terrorism in this sense is like the top of the iceberg. The other use of Internet, which currently constitutes the bulk of terrorist Internet use, is more passively. It is about spreading the word and communicating with peers, essentially the use that most people make of Internet and therefore exactly touching on the vulnerability of it. The second, passive, type of Internet usage will be my main concern in this thesis.

Although not all authors use the same terms or nomenclature, it is generally agreed upon that there are five core terrorist uses of the Internet (Conway 2005: 3-4). Other authors use more or less the same terms for similar matter¹².

1. Information provision¹³
2. Financing

¹¹ More information about social network analysis follows in the next chapter.

¹² Note that ‘cyber terrorism’ does not appear as a core use; this is really only a small fraction.)

¹³ Also dubbed ‘psychological warfare’

3. Networking
4. Recruitment¹⁴
5. Information gathering

Terrorist organizations are certainly benefiting a lot from Internet. An Al-Qaeda training manual captured in Afghanistan reads: ‘using public resources openly and without resorting to illegal means, it is possible to gather at least 80 per cent of information about the enemy’ (cited in Conway 2005: 14). It has been suggested that the Internet serves the very goal of terrorism exactly right. Terrorism, in the sense that Stohl defined it, is a theatre. It is a symbolic act or display. A rhetoric genre in the context of the World Wide Web. Indeed, publicity is an overriding consideration in planning terrorist acts (Weimann 2008: 72). However, only the huge displays of violence, the size of 9/11, Madrid, and maybe the assassination of Theo van Gogh, make it to the front pages worldwide. Therefore in the past, targets, location and timing had to be planned meticulously in order to pass the media threshold of newsworthiness, access, timetables and deadlines (ibid.). With Internet this media threshold is easily bypassed while an audience of tens or even hundreds of millions can be reached fairly easily. Moreover, ‘the Internet allows a person or group to appear to be larger or more important or threatening than they really are’ (Thomas 2003: 116).

And so as a symbolic act the phenomenon terrorism can then be analyzed like other media in terms of communication strategy:

Transmitter <i>(Terrorist)</i>	Message <i>(Bombing, ambush, etc.)</i>
Intended recipient <i>(Target)</i>	Feedback <i>(Reaction of target audience)</i>

Figure 1. *Communication strategy and terrorism*

Karber (1970, cited in Weimann 2008: 73) points to these four factors involved in a communication strategy. As conventional media make this consideration, why would not terrorists make it as well?

Several groups can be distinguished as targets or recipients of terrorist websites. Current and potential supporters, but also enemies, the international community and journalists. The target group may vary among

¹⁴ More problematic, only two authors, Weimann: 2004 and Thomas: 2003, include this category in their definition. There is, however, compelling evidence that indicates that Internet as a tool for recruitment has been widely used among terrorist organizations.

different sites. The goal of the website may indicate the audience for which it is meant. Some websites are available in English, pointing to an endeavour to reach the international community or media, whereas others have a more local appeal, keeping the audience up to date with recent activities and developments, or even selling items such as printed shirts, badges, flags, and video and audio cassettes. Also, slogans and rhetoric can often be more appealing to one of these target groups than others. So what's to be seen on these websites? I have already indicated that the first and foremost aspect of internet for most internet users, and likewise for terrorists, is the provision and sharing of information, either in the form of multimedia or in the form of text. Libraries of speeches, training manuals, extensive multimedia resources and links to other sites sharing similar beliefs may also be among the content (Qin et al. 2007: 72) Another important, interaction facilitating application that many terrorist organization maintain is a *forum*: an online discussion community, for which users may sign up. It is usually organized in topics and threads to ensure a focused conversation. Because it is not open to everyone a certain threshold of interest has to be passed. The organized nature of the forum makes it ideal to see what kind of discussions and opinions are present.

A commonly found element on those forums is texts providing an explanation or justification for the use of violence, with which terrorist groups are usually associated. This happens on radical Islamic forums all the time¹⁵. 'Terrorist rhetoric on the internet tries to present a mix of images and arguments in which the terrorists appear as victims forced to turn to violence to achieve their just goals in the face of a brutal, merciless enemy devoid of moral restraints' (Weimann 2008: 82). This technique of convincing the self that ethical standard norms do not apply in a particular context of violence is very common among the military. Using Albert Bandura's theory of selective moral disengagement, Weimann describes seven types of justification for a group's violent or even terroristic behaviour: displacement of responsibility, diffusion of responsibility, dehumanization of targets, euphemistic language, advantageous comparisons, distortion of sequence of events, and attribution of blame (2008: 79-80).

Also, e-mail, chatrooms, message boards, and online groups, are among the widely used online services that terrorist groups are currently and increasingly using. Without being aware 'Yahoo has become one of al-Qaeda's most significant ideological bases of operation' (Weimann 2008: 74), providing all these kinds of services for free. Also, training manuals and instruction videos on e.g. how to make explosives can be found on Internet sites. Thomas (2003: 113) lists a number of websites where such instructions can be found, contributing to what he calls 'cyberplanning': 'the digital

¹⁵ DWFP, see for example the search results of 'crusader' or 'mujahidin'

coordination of an integrated plan stretching across geographical boundaries that may or may not result in bloodshed' (ibid.). Here lies a huge potential for terrorists and various studies have revealed that internet played a very important role in the planning and coordination of events like for example 9/11, and various other, failed or 'successful', plots (see Katz & Devon 2007; Stohl 2007; Qin et al. 2005).

The influence of Internet is considerable. The Muslim Internet community, sometimes dubbed 'the virtual *ummah*', appeals to isolated individuals. It eases loneliness through connections with like-minded people. It leads them to spend more time with virtual community at the expense of social interaction with the direct social environment of such individuals, causing in turn further disembeddedness in society. The advantage for terrorists is that the virtual community no longer ties to any nation, thus fostering a priority of fighting against the far enemy (the West) rather than the near (Chen & Chung 2008: 1348). Chat rooms and forums tend to encourage extreme, abstract, but simple solutions, thus attracting most potential jihad recruits who are not Islamic scholars. Psychological studies have compared the strength of positive feelings that people develop toward each other online and offline. It seems that online feelings are stronger in almost every measurement (Sageman 2008: 114). Indeed, if anything, it seems that the true leader of the global salafi jihad is the collective online discourse of a dozen influential jihadi forums (ibid.).

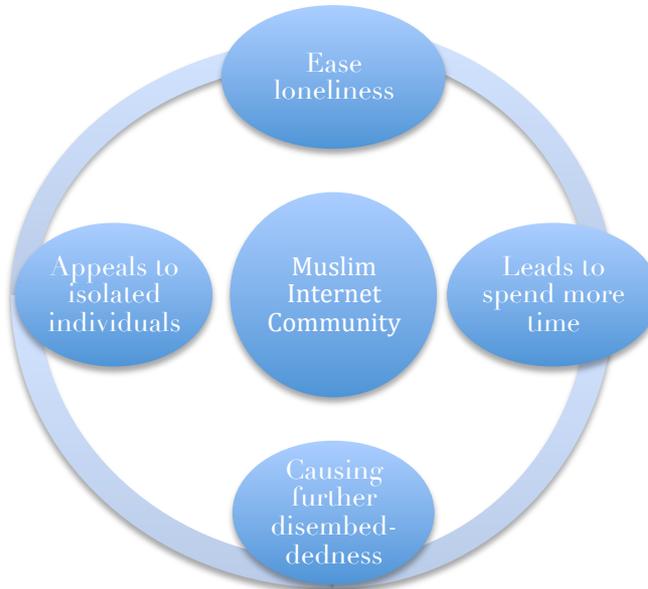


Figure 2. *A cycle of alienation from society under the influence of radical Islamic forums*

However, just interaction on the Internet is not enough to cause radicalization: devotion to jihad must also be fostered by an intense period of face-to-face interaction, according to Sageman (2008: 140). And so interaction in mosques might be of importance. The Al-Tawheed mosque in Amsterdam in the past repeatedly was associated with terrorists (Benschop 2005). However, though maybe a cliché, it should not be forgotten that

Islam, like most religions, preaches peace. Considering mosques to be centres of radicalization goes against the nature of the sacred house. Of the many mosques worldwide, only a fraction has a fundamentalist nature. Of those, only a handful support the global jihad. Despite being conservative institutions, mosques generally are not supportive of the global jihad, preferring submission to God's will rather than propaganda by deed. The socialization in and around the mosque builds and reinforces ideological commitment to the global jihad, which works as a process of bottom-up self selection. Adherence to the jihad, and therefore growth of the network is thus not a random process but one of preferential attachment, meaning that the probability that a new person will connect with any given person in the network is proportional to the number of his existing links with it (Sageman 2004: 139-140). Indeed, the more orthodox salafi mosques *are* places where extremists meet and network, and they do provide extremists with a large pool of recruitment, but they are not the primary sources of radicalization. Indeed, no one is converted by a website alone. It is the discussion with friends and family that contribute the most to radicalization. However, this discussion can take place online; it can be even more convincing that way. And so the question remains whether the idea of face-to-face interaction as being essential still resonates with the on-going rapid growth of the Internet and the development of terrorist use of it. The growing technical sophistication and content richness of terrorist websites (Qin et al. 2008: 83) make it ever more likely that face-to-face interaction is indeed obsolete in some instances.

Many of the authors referred to in the above section 'Internet', have avoided a too much technical approach. Remaining questions, therefore, might include: what is the size of this 'dark side' of the web? What major Internet technologies have been used on terrorist websites: what kind of interaction applications can be found? How advanced or sophisticated are these, and how do they facilitate a climate of radicalization? To gain an understanding, as well as to find hard evidence in order to avoid solely qualitative analysis while bypassing quantitative research, the technical aspect should be part of the analysis. Using even basic forum administrator's tools, it is possible to tell a great deal about the social structure of a virtual community. And so not only because of the nature of the subject demands it, but also because it can lead to a better sociological understanding of what is going on online, in-depth Internet research, I think is key to understanding what online radicalization is about.

I hope that in this chapter I have been able to make clear that the presence of some correlation between Internet and radicalization I think is undeniable. Further, I have demarcated the field that I want to research in further depth, by zooming in from terrorism to the global salafi jihad to radicalization to Internet. I have elaborated what kind of information terrorist organizations and extremists spread on the world wide web, and

what other uses terrorist organisations make of the Internet. I have also elaborated some ideas on radicalization and what it may lead to. One of the most influential authors on this field, Marc Sageman, has been criticized for failing to employ social network theory and associated automated methods to evidence his argument (Hoffman 2008: 87), and apply it for example to the present structure of Al-Qaeda's Internet component. This is a quite difficult step as it moves from social sciences toward informatics, but I think it is a necessary step. This step is an effort to synthesize the field of Conflict Studies with that of security Informatics. The Internet, in general, is characterized by its overload of information. This applies equally to the terrorism domain. There are a great number of possible extremist or terrorist sites, while only a few have a true dangerous potential. A possibility to cope with this is to sample a number of (representative) web sites, but mostly to stick to content analysis. Another preferable possibility, however, is to use a systematic methodology covering a much larger area. The next chapter will be about systematic web mining and web crawling techniques for researching the Internet.

3. Quantitative Methods for Research on the Internet

Since 9/11 there has been an increased interest in using information technologies for countering terrorism (Chen et al. 2008: 1349). Ironically, it were terrorist efforts to prevent authorities from tracking their web addresses and activities that have prompted several information services to monitor the websites of militant Islamic groups from the end of the 1990s (ibid.). The interest is evermore fed by the growing presence of extremist on the Internet, and has given rise to the academic discipline of terrorism or security informatics: ‘the application of advanced methodologies and information fusion and analysis to acquire, integrate, process, analyse and manage the diversity of terrorism-related information’ (Chen et al. 2008: XV), which had a considerable impact: ‘[t]he Internet has created a concrete bond between individuals and a virtual religious community’ (Sageman 2004: 159). Growing sophistication and interactivity of such web sites has effectively led to a ‘terrorist news network’, operating through their own Web sites and online forums (Qin et al. 2006: 71). Therefore the analysis with terrorist Internet use should inevitably have common ground with informatics, if you want to make sense out of the obscure tangle of data that is the net. To understand the networked structure and the tactical as well as strategic use of the web advanced methods are needed. In this chapter I will discuss a number of those methods that have eventually led to the creation of the Dark Web Forum Portal collection, that I have been granted the privilege of working with for my research. The Artificial Intelligence Lab at the University of Arizona particularly played an important role in the founding this collection of terrorist web sites and forums, which they have called the ‘Dark Web’. Many of the scholars that are cited in this chapter are linked to the Dark Web Terrorism research unit at this university¹⁶. I would like to explain how it was created and assembled, what kind of results it can produce, and finally how I intend to use it for the purpose of my research.

Terrorists usually operate in networks, either large or small ones (Sageman 2004: 137). Within these networks it is fairly easy to reach one another, as lines are short. Also online there are in fact many of these closely clustered networks. There are many with a few links, but only a few with many links. This applies in general for the whole Internet (Xu et al. 2006: 369); think for example of your own facebook-network. To be able make to such statements about the structure of online networks, advanced research methods are necessary. The immense size of the Internet makes it impossible to conduct research manually, which would literally take ages. And, in this particular case, besides the overload of information, the covert nature of terrorism and extremism makes it very hard to even locate such web sites. Jihad terrorism content on the web falls under the category of communicative contents and

¹⁶ Please see <http://ai.eller.arizona.edu/research/terror/>

so a quantitative analysis is critical for a study to be objective (Reid et al. 2005: 3-4).

Until quite recently the comprehensiveness of Internet researches was limited because scholars did not make use of automated technologies, but instead relied on observation and content analysis (Qin et al. 2006: 72-73, 74). Since a few years, intelligent automated systems have been developed to search the entire net for (potentially) terrorist web sites. The first step in the direction of efficiently analysing web site content is to capture relevant websites and store them in a repository. There are several ways to do this, and not all of them are equally efficient. The first is to start collecting a series of seed URLs¹⁷ that are manually given, and continuing with its direct neighbours by clicking and following links. This is meticulous, but also very time consuming. Another way is to have an automated program, a so-called WebCrawler¹⁸, search the Internet for relevant items. This is very efficient because it is fully automated, but it also tends to introduce a lot of noise – irrelevant items – into the collection (ibid.). A third and, according to many scholars, preferable option is to combine the two previous approaches into a new, semiautomatic one, where seed URLs are given, but neighbours and neighbours of neighbours are searched automatically (ibid.). This combines the accuracy of the manual approach with the efficiency of the automated approach.

As I said, in my research I have gratefully made use of the Dark Web Forum Portal, which has been created using semiautomatic data collection methods. When starting to assemble such a collection it is first important to define what you are looking for. The term ‘terrorist’ has very negative connotations and is politically loaded. People obviously never define themselves as terrorists. Therefore definitions by political authorities, NGOs, and academic studies were used. With the information from these institutions a terrorist lexicon was developed, including the name of the groups, sub-groups, leaders, jargon, etcetera, which could then be inserted in the major search engines. Some terrorist web sites could be identified directly via the before mentioned institutions. These web sites as well as the results returned by the major search engines served as seed URLs. With simple link analysis programs that follow both out- and in-link the collection was extended. Many websites that previously could not have been found or that were hidden with this method came to light. Language experts checked forum and other interactive content, and an advanced filtering system was created, both to prevent noise from entering the collection. Once identified,

¹⁷ With a ‘seed URL’ is meant one or several websites that have been given or found manually, to start the (semi-) automated research with.

¹⁸ A WebCrawler or Spider collects all links within a given lists of web sites, and adds to that list all new links that he finds. It is thus able to research the Web in an extremely efficient way.

the content is then downloaded. For forums, which usually are closed-off communities a program was designed which would apply for membership and then, once approved, collect the content (Zhou et al. 2006: 624-625). The above method, often called ‘web mining’, has resulted in a number of interesting articles about the Dark Web: the reverse side of the web used by terrorists.

First it may be interesting to look at the whole picture: the global level. What is the topology¹⁹ of the Dark Web? Web site interconnection relations provide a reasonably accurate representation of terrorist groups’ interorganizational structure (Xu et al. 2006: 368). Xu et al. found that there are three large components of extremist web site clusters: Middle Eastern, Latin American and US domestic (2006: 375). Also they concluded that each of the three components contained several websites that are not very large and share the same interests. It follows that they have more inter-site links and closer relationships (ibid.). Of the three large component networks, the Middle Eastern is the most active, which is indicated by the size and the high number of present connections between and within. The Middle Eastern network contains dense and strongly connected local clusters. Similarities in terrorist groups’ ideologies, beliefs, interests, and geographical closeness may cause their websites to frequently point to each other (ibid.).

Also on website level it produces interesting results. Websites of extremist or terrorist organizations have become increasingly more technically sophisticated (Qin et al. 2006: 83). In an extensive quantitative research, Qin et al. have been able to identify some key characteristics of terrorist web sites. Using a WebCrawler and a self-designed attribute system they conducted a systematic content analysis and a quantitative assessment of technical sophistication and content richness of terrorist web sites. This software is relatively easy available and also not too hard to build, for someone affiliated with informatics. They used this framework to compare Middle-Eastern terrorist websites with U.S. Government websites. All in all, Qin et al found that there is no significant difference between the technical sophistication of the Middle-Eastern terrorist websites and the U.S. Government websites (Qin et al 2007: 84).

Besides the result of this interesting comparison Qin et al. could draw some other conclusions of terrorist Internet use. They found that with increasing Internet surveillance, terrorists had to be more thoughtful and flexible in order to avoid discovery. And so for example they use free web servers, unsecured and poorly maintained commercial servers or proxy servers to

¹⁹ Topology is a branch of mathematics that is concerned with properties of objects and space under continuous deformation. It is used in network terms to refer to the consistency and structure of a network.

mask ownership (Qin et al 2007: 72). In 2006, Zhou et al. found with automatic web mining techniques that there are, within the US only, 110 extremist forums. Of these, 18 were hosted on extremist websites, 31 hosted on Google groups, 47 hosted on Yahoo groups, 9 hosted on MSN groups and 5 hosted on AOL groups. What is interesting to see is the high percentage (84 per cent) of forums hosted on public domains, which are easy both to start up and to abandon. The majority of the found files are dynamic files, mostly forum posts, pointing out that terrorists are relying on Internet-based communication tools (e.g. chatrooms, forums) to facilitate their daily communication, command and control, and even operation planning and coordination (Qin. et al: 2006: 82). In terms of files volume the vast proportion of the files were multimedia, especially video files, pointing at the heavy use of such files.

Indeed, as a research by Salem, Reid and Chen (2008) pointed out, videos are a very powerful tool to communicate messages quickly. Videos are a relatively cheap means to produce lots of information in little time. They provide persuasive action, vivid images and sounds that can highly appeal to the viewers' emotions, the importance of which I also have stressed in chapter 3. Just like terrorist forum content, terrorist videos have also been downloaded and stored. Most of the videos are quite amateurish: filmed in real-time, poor quality and simple. The 'violent attack video' was the most frequently identified (Salem, Reid and Chen 2008: 618), in which usually western military targets are attacked. Also scenes providing emotional and spiritual support by hugging and greeting one another and praying together are often seen. 'The videos convey messages that we believe are powerful enough to mobilize members, sympathizers, and even new recruits' (ibid.: 619). This all provokes a lot of discussion: in forum posts is often referred to videos and replies on videos on YouTube are almost a forum on itself. Moving on with this point are Bermingham et al., who interestingly argue that there is no division between information providing, but passive, websites and interactive means of communication such as forums at chat rooms (2008: 2). Indeed there are means that combine the two, such as YouTube. The more accessible YouTube is an earlier phase in radicalization, where forum users are more already made-up minds (ibid.). This may be so; it is however the latter phase of radicalization that I would like to focus on.

Another interesting aspect of the research by Bermingham et al. (2008) is the so-called sentiment analysis. They used automated 'lexical analysis' to determine what users were talking about, and 'sentiment analysis' to determine the users' opinions and attitudes. With a lexicon algorithm that assigns positivity or negativity to a word, in a particular context, they were able to define the sentiment of users toward certain terms. They summarized the 50 most frequently used terms into eight 'concepts' toward which user opinions were either very negative or very positive: 'America',

‘Christianity’, ‘Islam’, ‘Israel’, ‘Judaism’, ‘Mubarak’, ‘Palestine’, ‘al-Qaeda’ (Bermingham et al. 2008: 4). A ‘concept’ includes all variants of the word, for example ‘Jesus’, ‘cross’, ‘bible’, ‘Christian’, are all considered relevant terms with regard to ‘Christianity’ (ibid.). This way you can get a pretty good idea what kind of terms and keywords that may point toward radicalization, to look for.

I have highlighted the importance of (semi-) automated data collection methods for research on the Internet, which are indispensable for objective and well-informed research regarding online interaction. I have discussed some of the possibilities of those methods, including data mining and WebCrawlers. The application of these methods to the domain of extremist websites and forums has led to some important insights about the global structure of the Dark Web, as well as about the technical sophistication and richness of web sites. It turns out that terrorist web sites with regard to interactivity are well matched to U.S. government web sites. An example of this is the increasing amount of videos found on terrorist web sites: these days an important feature, spreading a powerful and inspiring message and subject of many online discussions. Finally I discussed how the use of certain keywords in online discussions may point to radicalization. My research is aimed at the use of online forums as interactive media of radicalization. I have already argued in chapter 3 what the role of forums can be in the process of radicalization. Another interesting aspect of forums is that all posts are preserved so that users may view them later and reply on them, unlike for example chat rooms, where content is not stored. And so forums are very rich in information, plus they are relatively accessible to outsiders (even WebCrawlers!) due to their size – a public forum typically has several thousands of users²⁰. However, with such an overload of information, it is easy to get lost. Also it can be very difficult to try to deduct social patterns and structures from such an enormous pool of data. Therefore in the next chapter I would like to discuss a number of methods that are very suitable for dealing with these amounts of data when you are trying to find social structures in large networks.

²⁰ Please see the Appendix 1 for detailed numbers.

4. Social Network Methods for Studying Large Sets of Data²¹

Social network methods are a relatively new way of interpreting large sets of data. There is nothing very new or unusual about social network data. Like other sets of data it is represented in matrices. Indeed, network data can be understood using conventional methods. In some ways, however, the methods differ from conventional data analysis techniques in social sciences, such as cross-sectional survey research. ‘Conventional’ data, so to say, usually can be shown in rectangular tables, with the rows representing the subjects or cases, and the columns representing scores, attributes, et cetera. You can then compare how similar or dissimilar actors are to each other across attributes of actors or by how these attributes are distributed among subjects. Network data, by contrast, usually consist of a square table of data. The rows of the table are the actors or subjects, and the columns are exactly the same set of actors and subjects. The table then describes the relationships between the actors. And so the major difference between conventional and network data ‘is that conventional data focuses on actors and attributes [while] network data focuses on actors and relations’ (Hanneman 2005: 3).

A possibility, still, is to look at these kinds of matrices the same way you would look at an actor-attribute table. Similar to working with attribute data, it is possible to see what actors have many relation-ties and what actors do not. Or in the case of weighted or valued relations, the amount of connections that an actor has. This is a useful way to look at data, it shows to what extent actors are ‘embedded’ in the network as a whole. It is also possible to compare one actor’s ‘embeddedness’ with that of others. Also the centrality, power, and prominence can be deduced from how a node is placed within its neighbourhood. This is called social network analysis on micro-level. The other way of looking at the network data is more holistically, for example by looking at the density of the network or the reciprocity of relations. What is important in this view is how certain roles or positions are distributed among the actors. We can thus get an overall picture of a whole network. It is then possible to look for example at the density of a network. This can tell us something about the position of one individual as well. If a network X is a very dense network, but actor A has only a few connections, then clearly actor A has a problem with connecting or does not want to be connected or others do not want to connect to A.

Social network analysis is very much about switching between macro- and micro-perspective. The underlying idea is that all properties, social

²¹ This explanatory chapter was written with the help of Robert A. Hanneman’s handbook ‘Introduction to Social Network Methods’, 2005.

positions, or roles of an actor are inherently relational. For example, it takes a wife to know that someone is a husband, or it takes a capitalist to define a labourer. Patterns of regularity then define the boundaries of a social position (Hanneman 2005: 196). Again: what's important is not an actor's attributes, but his relations. The core elements of social network analysis therefore are nodes and edges. Nodes represent actors, and edges relations. So far this seems pretty straightforward, but you can imagine that this approach has important consequences for setting up research, sampling and analyzing data. While very common in social sciences, a network analyst on the contrary would rarely use sampling in his work, but rather focus on the whole. There are several methods how nodes and ties within a network are collected and built up into a matrix or graph.

First, a full network method would require that you collect information about each actor's ties with all other actors. In fact, this is like taking a census. The data collection technique is quantitative: the aim is to collect for example all data on oil shipments between all countries. This data collection technique guarantees the most powerful analyses and descriptions of social structure. However, it is not too hard to imagine that this method is expensive, time-consuming and difficult to collect. For example, in the case of oil shipments, it would be doable, but imagine that you would have to ask all students at a university to identify their friends, or count the total amount of traffic between two cities. With bigger populations, it usually becomes much more difficult to collect and analyze data. A second method, therefore, is to start with a small focal group and collect their ties to others. From the others, those that were not included in the primary list are tracked down and asked for their ties, and so on. This process continues until the decision is made to stop (due to the limited availability of time or money) or new additions are very marginal to the group that is the focus of the study. This method is called the snowball method. Internet search engines work through websites in a similar way, and the Dark Web Forum collection has also been assembled with this method²². A disadvantage is that isolated points are overlooked and therefore the measure of connectedness of a network is easily overstated. It is however suitable for, among other types of networks, business and family networks, and for capturing the elite within a larger group. It is worth noting that no single approach is the 'best', but rather that the suitability of the data collection technique is dependent on the purpose of the researcher.

Networks are usually displayed in a graph. A graph or *sociogram* is composed of nodes that are connected by edges, which represent actors and their relations amongst each other. There are many ways to display graphs. Different algorithms have been developed to display a network, putting different emphasis on different aspects of a network. Those tools are very

²² See chapter 4.

helpful for dealing with the complexity of larger networks. Concealing or emphasizing some parts of the total can be a big help in gaining a better understanding of the overall picture. They can help you seeing patterns and connections that you would otherwise not have seen. Indeed, a picture may say more than a thousand words.

Data can be arranged in a number of ways. The simplest way is in an undirected and binary graph. In such a graph, a connection exists or it does not. This works for some purposes, for example when representing one's facebook-friends, who are either connected or they are not. A connection on facebook is always reciprocal, as people have to confirm their friendships. Therefore such a graph works for this purpose. However, binary, undirected graphs often are too simple. When analyzing for example money or information flows, ties should at least be directed, that is, they should become arrows. Information flows are not necessary reciprocal and therefore an undirected graph would not accurately represent the reality. A directed binary graph would make sense, in such a case. Information flow is difficult to count or express in numbers. However, for money flows a valued directed tie would more accurately represent the situation.

Although the basic idea is simple, the further elaboration of the concept of social network analysis can be quite complex. A major part of social network analysis is managing an extremely large amount of data. Even with small or modestly sized networks this can easily become very complex and confusing. Therefore the use of a number of mathematic tools is indispensable. Usually the network is represented in a square matrix, called an adjacency matrix. This may seem a bit too formal, but it forces you to work with your data systematic and complete. Another important feature of social network analysis is to make use of computers. Representing network data in either a matrix or a graph enables you to apply computer technology to your analysis. Building up and analyzing larger networks can take years (literally!) if you would have to do it by hand, while a computer does it in a few minutes. Analyzing network matrices is work that is repetitive and boring exactly the sort of work computers do very well, but humans do not. The social network approach therefore particularly makes sense when applying it to an extremely large set of data. Finally, following the strict rules of matrix and graph processing helps you to display everything clearly, but also may lead you to see patterns that you would otherwise have overlooked.

How then to analyze a network? What can we tell from a graph or matrix? Even from the very simplest of network structures there is a great deal that you can deduct both about individuals and about the population as a whole. I'll walk through some very basic properties of actors and networks. First, the focus can be on the immediate neighbourhood of an actor. An actor's *degree* is the number of connections that an actor has; this can be further distinguished for directed networks into in-degree (connections coming in)

and out-degree (connections going out). This tells the extent to which an actor is connected or constrained to his environment, and so it is possible to get an idea of the opportunity structure of an actor: what are his opportunities to receive and disseminate information? Another indicator regarding this question, but then on an overall level, is the *density* of a network. The density is the proportion of all present ties from ties that could theoretically have been present. Again this can be applied to both directed and undirected networks. Degree and density are useful numbers for explaining social behaviour as they show, respectively, how connected an actor is to his environment and how connected the whole network is in general. Finally, the diameter, the largest ‘shortest pathway’ in the network, can be a good indicator for the size of the network. For example, a diameter of 3 shows that a network is very small, in which information can reach all actors fairly easily, while a diameter of 11 is extremely large or shows that the network is extremely fragmented. All of the above numbers refer to the connectedness or embeddedness of an actor and of the whole network. For different purposes, different numbers might be relevant. There is no single right way. These methods help in quickly gaining a basic grasp of what is going on in a complex network.

The next step to get a more comprehensive picture of the structure of the network is measure the extent to which a network deviates from random connection, or from archetypes of hierarchic networks (Hanneman 2005: 145-146). We can thus deduct connections of hierarchy from the network, and we are getting close to the concept of deducting social roles from the position of an actor in the network. Therefore, we first focus on the direct neighbourhood of one actor; that is, his *ego-network*. Usually an ego-network consists of one actor and ties to his direct environment, and the ties among his neighbours. ‘The fundamental idea here is that the ways in which individuals are attached to macro-structures is often by way of their local connections’ (Hanneman 2005: 143). Access, opportunities and constraints are in most cases provided by the direct local environment. A particular interesting statistic, is called ‘*betweenness*’, and basically means the extent to which ego is between two others. This says a lot about ego’s supposed prominence in the network, for example the extent to which alters need ego to reach each other with information. A certain ego may have a lot of ties to other actors. Those actors, however, may be rather isolated. Ego is therefore a prominent actor, but only in a very local area.

An important aspect of social structures, many social scientist would agree, is power. Power is as important as it is complex, and there is little agreement on what power is and how to analyze the phenomenon. Within the network approach, however, it is emphasized that power is relational – that is, where there is power there is also weakness. Power is not an attribute of an individual, but rather it derives from the relations with others. In other words, ‘[p]ower arises from occupying advantageous positions in networks of

relations' (Hanneman 2005: 168). With 'advantageous' can be meant a high degree or high betweenness, or other measures. And so actors with a high number of incoming ties can be thought of as prominent, because many actor want to establish ties with them. On the other hand, you may think of an actor with a high out-degree as being an influential player. Similar statements can be made about betweenness, or other measures of centrality. A network may have a high or low level of power as a result of patterns of connections. With this kind of analysis it is important to often switch from micro- to macro-perspective. However, often researchers are most interested in the social units that lie between those two levels. Sageman (2008: 23-24) likes to speak of the meso-level. I will now turn to those sub-structures.

Whereas centrality and power are a top-down approach for unravelling the structure of a network, finding sub-structures often, though not always, is more a bottom-up approach: building up a network out of smaller and larger interconnected components. Another, indeed more top-down look at it is looking at the substructures as a very dense, and somewhat separated area of a graph. Both approaches may be feasible for certain purposes. For locating sub-structures, most algorithms simplify the network, leaving aside the distinction between directed and undirected ties. Usually strong ties are used, that is, only ties that are reciprocated count as ties, others are ignored. It is then possible to distil a number of types of sub-structures. The most ideal sub-structure is a *clique*. 'A clique is the maximum number of actors who have all possible ties present among themselves' (Hanneman 2005: 175). Groupings in a network have a great explanatory value for behaviour of actors. It shows how information can travel across the network, but also how and what kind of conflicts can take place. Some actors may be rather isolated in terms of cliques, or have a position in which they can act as a bridge between groups and maybe even exploiting such an advantageous position.

However, often structures such as a clique in the rather strict definition above may be hard to find, or there are so few within the network that it does not make much sense. It is very likely that actors will consider themselves a member of a group if they know a certain number of other members, but not all, or not even most of them. Or they have not as strong ties to group members as others have. This has some interesting implication for identity questions regarding group membership. It can also be interesting to look not at inclusion, but exclusion. On a rather moderate forum a certain clique may exclude themselves from the rest of the users. Such a clique may be called a *component*: sub-groups that are well connected within, but disconnected between other groups or actors. These kinds of areas in the graph are denser than the graph as a whole. Indeed, '[d]ifferent definitions of what a clique is can give rather different pictures of the same reality' (Hanneman 2005: 193-194). Therefore another useful statistic is the

clustering coefficient, which can give an overall indication of the extent to which the network tends to cluster together.

With a structure of sub-groups, cliques, cores, or components established, it is now possible to distinguish between in-group and out-group members, this is often called to make ‘partitions’. Group membership may also be assigned manually, should the researcher be well aware of the situation. It is then interesting to look at pairs of actors that are not directly connected, but that are connected through another actor. In several ways, that actor may benefit from his social position by brokering. The idea of brokerage is that ego is the “go-between” for two actors – that ego falls on the path between them (Hanneman 2005: 135). The degree of potential for brokerage depends on the before mentioned degree of betweenness. The way an actor is embedded in his neighbourhood is very useful in understanding power, influence, and dependency effects. The supposed inclination of every actor is that of a rational individual actor who may be attempting to maximize profit or advantage by modifying the way in which they are embedded. The membership of a group then gives an idea of the type of social brokerage role ego can have. Brokerage is also very much linked to leadership roles. The below table lists the different types of social brokerage roles depending on group membership. The middle actor is our focal actor, the direction of relations is from left to right.

Figure 2. *Social brokerage roles depending on group membership*



Group membership			Social Brokerage Role
A	B	C	
In	In	In	Coordinator
In	Out	In	Consultant
Out	In	In	Gatekeeper
In	In	Out	Representative
In 1	Out	In 2	Liaison

I have discussed the basic principles of social network analysis. Such methods are essential for making sense of the interaction on a forum with sometimes over 1,000 posts per day. By focusing on relations between actors, instead of comparing individual actor’s attributes, you get a much more organized picture of, in this case, social traffic and flow of information – but any kind of traffic could be filled in there. Basic functions of analysis include partitioning the whole into cliques, components, or other kinds of subgroups, as well as ranking individual actors in terms of degree or betweenness. Furthermore it is possible to look at the interaction between those two social units – cliques and prominent actors. For the purpose of

this research, I think the interaction between cliques as a whole and actors in advantageous positions, that they may exploit as social brokers, is of particular significance. Drawing from stories like that of Mohammed B. and work of scholars like Sageman (2008) it is my hypothesis that radicalization through Internet forums typically occurs clique-wise. Small groups move along the path towards radical Islam pushed by influential actors. In network terms, radicalization can be seen as exchange of information and asserting influence deriving from power, centrality, and prominence, on others. Interaction on forums is all about networking and developing relations. Therefore the social network approach is an excellent way of exploring social structures within terrorist and extremist forums.

5. Methodology

In this chapter, I will explain how I intend to make sense of the extremely large Dark Web Forum Portal database²³. Let me just quickly recapture what I have argued so far. Online jihadi forums, it is widely agreed upon, play an important role in the process of radicalization of young Muslims (Bermingham et al. 2008; Sageman 2008; Qin et al. 2007: 73). Whether forums contribute to mobilization or merely to radicalization remains to be seen, but the influence of them is recognized. There is plenty of data: an extensive network of websites has been established by numerous terrorist groups, a network capable of reaching a diversity of target groups ranging from potential recruits to western-oriented media and many more. Tools have been designed to research the entire Internet for potentially extremist or terrorist groups using WebCrawler technologies. The findings of more than a decade of semi-automated research have been organized and stored in the Dark Web Forum Portal, a portal created by the University of Arizona, an important centre for research on ‘internet and terrorism’. All posts on all terrorist/extremist forums worldwide are accessible through that portal. With social network analysis tools and software it is possible to make sense out of this overwhelming pool of data. For example, it is possible to identify cliques, influential players and a numbers of other attributes of forum users. So in the next step, using a structural analysis and social network methods I want to investigate the importance of cliques in the process of radicalization through online forums as well as look at patterns of interaction between cliques and prominent actors. I hope to be able to show that radicalization is a process that occurs in cliques: small groups of close friends who move stepwise closer toward extremism. Having identified the cliques with this analysis I also hope to reveal a major catalyst for radicalization: the power of influential actors on the process of clique-wise radicalization.

In the Dark Web Forum Portal it is possible to search for a (limited) number of keywords. It then identifies on which forums the keyword has been used, how often, by whom, and on which thread²⁴. It is also possible to display the relations between different users in a social graph. Social graphing and social network analysis are particularly helpful tools when researching the Internet, forums, and discourse in general. I will work with binary data – a connection exists or it does not – and valued data, where the amount of forum posts on the same thread also counts. When identifying cliques as tight social groups it might be necessary to put up a certain value-threshold for recognizing cliques, to ensure an understanding of cliques as tight social

²³ I will henceforth refer to the Dark Web Forum Portal using the abbreviation DWFP, followed by a thread number, or another part of the portal. Please see the Bibliography for the full reference and URL.

²⁴ A ‘thread’ or ‘topic’ is a chain of posts on a forum around a particular subject, issue, theme, or event.

units. The definition of a clique in network terms is very strict, in social terms it is more ambiguous. The algorithm would recognize a clique already when all members would have posted only once on the same thread. In social terms, that is a very thin base for a clique. Therefore I decided to use another, less rigid partition called *modularity*. Modularity tries to optimize the quality of a division of the whole network into communities – a good quality division being one with dense internal connections but sparse connections with out-group users. It strives to create an optimal scale-free network. Although the method is different, the results tend to be much the same. Plus, they are represented in a much more convenient way. I have also looked at the amount of interaction between clique members. The number of cliques that could socially be recognized as such dropped using this criterion.

My plan is to try with several different keywords to find forums and threads where radicalization is taking place. When I have found concrete examples of radicalization I want to map the social structure of the forum.

That is the idea. Now how to work that out concretely? First it is important to define what the nodes and edges would mean in this particular case. The nodes here represent persons, or more specifically forum users. Theoretically one person could have multiple accounts and therefore counts as two users. So a user is always a person, but a person is not always one user. Secondly, the edges represent a forum post on the same thread. It means a flow of information. A thread is like a conversation, and therefore it is likely that users, once they have posted on a certain thread, will check back regularly to see if other people have replied on their post or to watch the development of the discussion. Therefore it would be valid to say that there really is interaction on forums and similarly to regard users who are repeatedly active on the same threads as (virtual) ‘friends’. Plus, many more people have only read the post(s), but not replied on them. And so the number of replies on that thread can give us an idea of the impact of a certain message.

The first selection criterion is keywords. When doing research on the Internet this is a very common selection criterion. We use it all the time when searching the Internet with Google. To use keywords for selecting a part of a discourse is not unusual either. With explorative research, such as mine, it is a viable method. Some remarks, however, should be made. A keyword search may not discover all relevant results, or in this case forum threads. Due to the covert nature of terrorist forums people may choose to defer from using words that point toward terrorism in order to avoid detection. Instead they may use metaphors or other references. Keyword search overlooks this. Also, keywords are not the same as subjects; the fact that a keyword is mentioned in a discussion does not necessarily mean that the keyword is the main topic of discussion. In general, the more the

keyword deviates from ordinary everyday lingo, the more significant the results will be. Keyword search can be rigid and sometimes lacks flexibility. Plus, the significance of the keyword varies. However, it is the only way if you want to make sense of discourse on forums with over thousands of users. But I believe that keyword search is a good tool to work with sets of data this large, where browsing everything simply is not an option. I will try using several different keywords that relate to radicalization. But we should not forget that the Dark Web research team at Arizona University already has done a lot of the selection work. The forums they selected have been acknowledged to be centres of radicalization.

Now radicalization lingo often centres around violations against other Muslims (Sageman 2008: 72; Bermingham et al. 2008: 3-4). Therefore keywords relating to Christianity and Judaism – in extremists’ eyes the most important enemies of Islam – are very likely to point to radicalization, as well as ‘jihad’, although it is a much more ambiguous term. Furthermore, a term such as ‘crusader’ – a hate term for Christians ‘invading’ Muslim lands – very likely points to radicalizing discussions (Reid et al. 2005: 4), and tends to be bound to a geographical region, the Middle East²⁵. ‘Mujahidin’, also is an often-mentioned word, mainly in news messages, but also in some longer discussions. I thus hope to identify a number of forums and threads where radicalization is taking place.

The Dark Web Forum portal scans forum posts from the last 10 years or so on 29 forums worldwide that are considered terroristic. However worldwide the Internet may be, these forums tend to be connected to a specific region. Relatedly, languages also differ. Arabic, English, French, German and Russian are key languages among forums that have been recognized as being terrorist. Religious perspectives also differ, but most of them are in the salafi corner²⁶. I have decided to select one forum that is based in the Middle East, al-Faloja, and one that is based in Europe, As-Ansar, for research in further detail.

I can then download a dataset related to one or more specific keyword(s). That is why the search terms have to be considered well. I have to report that I encountered some problems working with the Dark Web Forum Portal. The module that enabled me to download a workable dataset is broken and there were neither funding nor personnel to fix it. I could still access the forums and read the messages though, but the online module that prepares data into a format that is compatible with the analysis software is malfunctioning. I have solved this problem by downloading the raw text messages from the forum and parse them into a workable format. For this a

²⁵ DWFP, ‘crusader’ cross forum search results.

²⁶ See Appendix 1.

program was designed²⁷ to turn those messages into a simple adjacency matrix, which identified nodes as user ID, and edges as thread ID. Those ID's are unique numbers that the Dark Web Forum Portal assigns to each user and each thread. The program can then easily identify which users are on mutual threads, that is, which nodes are connected to each other. Also, the program takes notice of how often users have posted on mutual threads, so as to assign weight to the edges: thicker edges mean more mutual threads between two users. In order to limit computation time I had to preselect the latest 200 users mentioning a specific keyword. With the smaller As-Ansar the message limit is not a problem at all, as with all search terms all messages are included, from the period December 2008 – January 2010. With the 15 times larger Al-Faloja I cannot use all data, but only 200 posts. However, for all search terms that I intend to use, the limited data almost covers the period when the forum was most active, between December 2008 and December 2009, and so I can take those results for representative. I will process the downloaded data with Gephi 0.8 alpha²⁸.

I have chosen to work with symmetrized data, that is, I left out the direction of the edges. In some contexts it can be useful, but on a forum direction of messages is very much arbitrary, since there is no such thing as a message exclusively directed at one user, as with e-mail, for example. Quite the opposite: all users are able to read everything that has been posted on a thread. Therefore the reach of a message should not be underestimated. What I have been looking at is online ‘conversations’, but only what has been written. Much more users may have read the contents of a thread, but not replied on it. The forum that I have investigated the most is the Al-Faloja forum. In its peak days more than 1000 messages were posted each day²⁹.

I found that the bulk of the threads are extremely short – only one message, the ‘newsflash’ kind-of thread I have described earlier. Usually these messages discuss actualities – although once I stumbled across a one-message thread with an extremely negative review of the book ‘Leaderless Jihad’ by Marc Sageman that I have discussed in chapter 3!³⁰ The one-messages threads do cause a lot of noise to enter the social network. In the samples I took every time more than half of the nodes were one-message threads, in other words, with a degree lower than one. Because there is no proof of interaction on those threads they are not really of use here, and so I have filtered them out. In the following chapter, unless stated otherwise, all

²⁷ My thanks go out to my good friend Sascha Ramondt, who was kind enough to design this program.

²⁸ Gephi Graph Visualization and Manipulation Software. Please see also the Bibliography.

²⁹ DWFP, Al-Faloja Statistics

³⁰ DWFP, Al-Faloja thread #75143

statistics are taken from the filtered nodes – the nodes with a degree larger than 1.

With the algorithms that I explained in the previous chapter I want to identify the cliques. In this case, a clique would consist of a number of users that regularly have had contact on the same threads where a particular keyword was used multiple times. The more often they posted a message on the thread, the firmer their bonds are represented. As I mentioned before, I want to make a certain threshold level for cliques to be recognized as such. For me, if all group members have posted only once or twice on the same thread, that would be too thin a base to call it a clique. I do not want to use a fixed threshold, but assess each instance. It is also possible that I am not able to find any such cliques at all.

My next step is to look at who are the most powerful, central, prominent and influential actors on the forum. I will use different definitions and algorithms to define them. The number of posts certainly is a good indicator of prominence. But power and centrality in this case not only depend on the amount of output, but also on the number of people that an actor may reach. An actor with a low degree centrality but a high number of strong ties can easily be a respected person, because he did not invest much timing in leaving many posts, but has many connections nonetheless. Clearly, people want to be in contact with such a person. An actor's betweenness³¹ may also indicate that he is in an advantageous position to influence others. Such an actor may, for example, serve as a mediator between Al-Qaeda or other global salafi jihad-affiliated organisations and the ordinary forum user. Another definition that I have found useful, especially in the context of the Internet, is PageRank. It is used by Google to rank web pages³²; but it works the same with nodes in a network. It ranks them according to how often a random user following links will reach the node. This gives more prominence to users with more posts, which makes sense, because it is more likely that their messages will be read, but also implies that positioning is very important. Furthermore, I will also look at brokerage roles, as described in the previous chapter, to see more elaborately in what kind of social position they are. Because I will symmetrize the data the number of brokerage roles will also be limited to 'coordinator' and 'consultant'. Direction of ties is not so relevant with those, but rather group membership. Therefore I find those terms the most applicable to the social reality of a forum.

³¹ The extent to which an actor is 'between' two others. Being in between is an advantageous position that can be exploited as an important source of power and influence. See also chapter 5.

³² But it is named after its founder, Larry Page, also co-founder of Google.

With (hopefully) a number of cliques and powerful actors recognized, I hope to find patterns of regular interaction between them. That is my actual hypothesis: that there are patterns of interaction on forums between cliques and powerful individuals, which encourage (further) radicalization among cliques. My research is explorative and aimed at finding patterns of social interaction at meso-level. I cannot deliver hard evidence that radicalization is taking place; I can, however, analyse how the social climate toward radicalization is by analysing the social structures on the forum. However I need to check back the texts of forum posts to see if it is still radical discourse that I am theorizing about – and not last week’s football game, for example. I choose to use social network analysis, and not content analysis, to find social patterns. Therefore I will randomly check the discourse, but that will not be my main analysis technique.

I have first summarized the focus of my research: the social process of radicalization within cliques, encouraged by influential outside actors. My most important source of data is the Dark Web Forum Portal, created and maintained by the Artificial Intelligence Lab of the University of Arizona. Using different keywords that are specifically related to radicalization I will search through the collection. From this huge collection I have selected a Middle Eastern-based and a European-based forum for further research. The text data that I downloaded from the forum have been parsed into a dataset that is compatible with Gephi, a social network analysis and visualization software. The data will be symmetrized – because direction in forum interaction is very much arbitrary – and the degree range will be filtered to exclude noise. Using basic functions of Gephi I can identify cliques, but also prominent figures, for which I will use different definitions. After having identified those, I should be able to see whether there is interaction between them. I will randomly check back the posts on the forum to ensure that I am still theorizing about radicalizing discourse. In the next chapter I will discuss my findings in detail.

6. Findings³³

In this chapter on my findings I would first like to give a general impression of the discourse I found using several keywords that likely point toward radicalizing discussions. Then I will discuss my findings on the social structures within the forum. Because the system works that way, I have grouped my findings per forum per keyword. After that I will discuss some more general social patterns that I have found. I have decided to focus my research on two forums. First, the Al-Faloja forum, since it is ‘one of the most popular and credible among jihadists’³⁴ worldwide. So the content is very distinctly radical and possibly mobilizing, whereas on other forums sometimes only sections are said to be radical. Second, the English branch of the Ansar al-Jihad network, because it ‘is popular with western jihadists’³⁵, and I think radicalization of Muslim youth in western countries is very interesting. Also, the Ansar forum is much smaller – it has 15 times less users than Al-Faloja – so different patterns of interaction might be visible, when compared to the larger Al-Faloja forum. Both forums are dead by intervention. Also different ways of analysis are possible: on the Al-Faloja forum it is possible to go into detail about postings due to the smaller size, while with the Al-Faloja forum I can look at the structures at a larger scale.

Discourse

I first tried searching on the term ‘crusader’, because the term is very much unambiguously directed against foreign western powers and military, which are viewed by extremists as primary evildoers and enemies in Muslim territory. On all forums the term is used. On the Al-Boraq forum the term is used significantly more often than on others. However, the forum is an affiliation of the Jihad and Reform Front, an umbrella organization consisting of the Islamic Army in Iraq and other groups. An increased mentioning of the term ‘crusader’ by a movement operating in Iraq confirms the commitment to the Iraqi insurgency. The search term ‘Judaism’ generally returned negative to extremely negative messages, very often related to the Al-Aqsa Mosque in Jerusalem. Violence and injustice, or even fear for those, is an important theme of discussion on the Al-Faloja forum³⁶. Further, untrue news messages are found, attempted to sow hatred and fear, but also calls to actively hack Israeli web sites³⁷. Search terms ‘mujahidin’ and ‘jihad’ returned many short news messages, sometimes followed by a

³³ I could not have written this chapter without the help of my friend Sascha Ramondt, who parsed all raw data into workable files. My thanks.

³⁴ DWFP Home.

³⁵ Ibid.

³⁶ DWFP, Al-Faloja threads #43883, #75252, #89758 (among many others)

³⁷ DWFP, Al-Faloja threads #40761, #49727, #39855, #39795.

discussion of approval and praise, but also a lot of links to YouTube and files that I assume mostly are videos or images³⁸.

Although it is hard to generalize when talking about 3,364 posts, based on the samples I took, the pattern I deduced from the discourse was the following. On the Al-Faloja forum the searching for the term ‘crusader’ mainly returned ‘newsflash’-like messages. The messages there are mostly related to the military conflict in Iraq, and usually are about losses of coalition forces, at the hand of the mujahedeen. It is uncertain whether the messages are posted by the perpetrators or by thirds, as many messages are written in the first person plural. The forum contained the same messages often twice: in English and Arabic. The messages in English apparently are meant for outsiders to read, but are often without any reply on them, while the Arabic section is much more alive, on the same newsflash-messages there are mostly between 10 and 15 reactions. Still the larger portion of such messages is without any reply. The reactions are mostly very much approving of the violence against the westerners. In religiously tinted wordings the perpetrators are praised and blessed.

Table 1. *Mentioning of the term ‘crusader’ on four Arabic and two English forums.*

Forum Name ³⁹	Mentioning of the term ‘crusader’ or ‘صل ي بي’ ⁴⁰	Total number of messages on forum ⁴¹	Percentage
Al Boraq	3,364	270,048	1.246%
Al Faloja	1,485	564,159	0.263%
Ana al-Muslim	1,754	1,603,518	0.109%
Shamikh al-Islam	1,344	299,956	0.448%
Al-Mujahiden Network	481	165,510	0.291%
Ansar al-Jihad Network [EN]	64	29,492	0.217%
Ummah [EN]	34	1,357,525	0.003%

What struck me immediately was that only a very small portion of all forum users was actually engaged in discussions where the radicalization-apt keywords were mentioned. The majority of the messages, in that sense, are noise. However, before you can generalize with social network data, you need to have a substantial set of data, that would be, 100 or more relevant

³⁸ DWFP, see for example Ansar1 threads #4618, #9852, #963 or Al-Faloja threads #7076, #55432, #5200.

³⁹ For this table I have selected the five Arabic and two English forums where the word was most used.

⁴⁰ DWFP, ‘crusader’ cross forum search results.

⁴¹ See Appendix 1.

nodes; otherwise it gets too speculative. I hoped rather to get an overload of information, but instead I have to mind over-interpretation.

‘Crusader’

The graph of users and threads in which the word ‘crusader’ was used on the Al-Faloja forum showed some interesting structures. The graph density is 0.146, in other words, ~15 per cent of all possible ties are present⁴², a medium high density for a graph this size. The clustering coefficient is

0.879, which indicates that there is a relatively high tendency of clustering together.

The diameter of the network is 4, meaning that information can reach everyone in the network within four steps. There is one group with a medium interaction that is, a medium number of mutual threads

among each other (group purple) and a group with much interaction on a lot of threads (group red). In the latter group there is one leading figure with stronger ties to the other group members than they have amongst themselves.

This user, #13534 (central in group red), also has strong ties with at least three other out-group prominent forum users, including user #11877, who has the highest degree and highest betweenness score, and also the

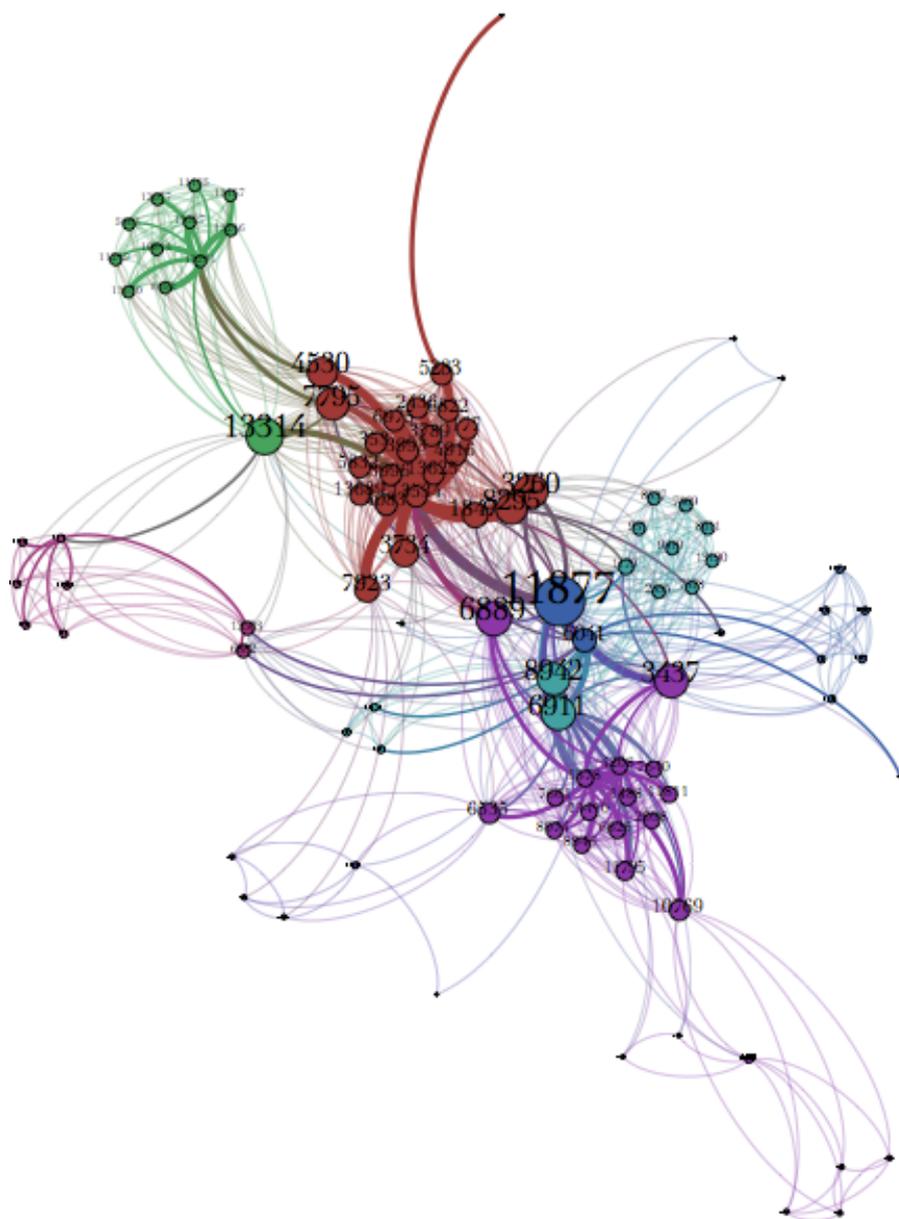


Figure 3. Graph of the filtered users of the Al-Faloja forum mentioning the word ‘crusader’. The size of the nodes indicates their degree; the thickness of the edges indicates the number of mutual threads between two users.

⁴² Please see also the Appendix 2

sixth most posts on the entire forum. 13543's degree, however, is not higher than that of in-group users. Without too much time-investment by leaving a lot of posts on the forum, 13543 has a loyal group of followers and is in contact with several very prominent users. This means that user #13543 is probably a much-wanted contact, both inside and outside the group, and thus in a position to be a possible source of radicalization. Considering brokerage, I can distinguish two types: a coordinating and a consulting role. The former applies to the actors within groups with no significantly higher degree than others but with more and stronger ties to other group members than they have among themselves – examples are user #13534 and the central user in group emerald. The latter applies to users with a high degree that on this representation are placed somewhat outside the groups – such as user #11877, but also #6889, #3437, and others. Density in the network is high and the diameter is low, which indicates that information can pass relatively fast throughout the network. This gives more prominence to consultants compared to coordinators. Given the high occurrence of coordinators in this network, the network is quite prone to radicalization.

On the Ansar forum, the term 'crusader' is mentioned by 43 users on only 64 threads. I find that too thin a base to generalise about the social structures. The users of this forum are generally based in the UK, and apparently do not hold such negative opinions on western interventions.

'Judaism'

Continuing on the Al-Faloja forum, when searching for 'Judaism'⁴³, there also is a lot of noise, but also at least one clearly demarcated clique (group violet), of which the members have been active on quite some mutual threads, enough to be able to call it interaction. User #6911, a prominent figure being the user with the most posts on the Al-Faloja forum, has a very strong tie with one of the members (#7065) of the aforementioned violet clique. It is possible that 7065 has a leading role within the clique and that he was being influenced by 6911. 7065 clearly has some radical ideas, on the forum calling for a nuclear jihad and expressing that he has links with the Islamic Army of Gaza⁴⁴. The density of the graph is fairly low (0.072), one of the lowest in the samples I took. The average degree is also considerably lower, only half than with the keyword 'crusader'. This means that a fewer number of people are involved in discussion and that threads are rather short conversations, they threads are shorter, they contain less messages and more fragmented. Also, degree and betweenness are more unequally distributed. The nodes in the graph are sized according to PageRank, so the size of the nodes indicates the probability that a random user following links will read the posts of that user. Given the global content of posts related to 'Judaism', which I discussed earlier in this chapter, users browsing these

⁴³ Which includes morphological variants, such as 'jew', 'jewish', etc.

⁴⁴ DWFP, Al-Faloja thread #403640

posts are probably prone to anti-Semitism. The size of the nodes gives an indication of the contribution of each of the users to that sentiment on the forum, by indicating the probability that such expressions are read.

A handful of actors are in a position to influence a vast portion of all users active on threads about 'Judaism'. However, as the total users is much smaller than for example with 'crusader', that influence is limited. The brokerage pattern that I found mainly is that of outsiders having strong ties with one or a few group members consulting brokerage roles, that is. The low density and high diameter of this network indicate that information is not circling around too fast. The power of consultants decreases because of this structure. Therefore the role of in-group coordinators should be high if any significant amount of interaction is to take place. However, within the few clusters of this network degrees generally are low, pointing to limited interaction. Given the relative large number of users with very few posts, only a few members are active on this subject. Hence radicalization on this topic probably is not occurring as much as with 'crusader' on the Al-Faloja forum.

The number of users mentioning 'Judaism' on the English Ansar forum again was too small to be able to generalize from it.

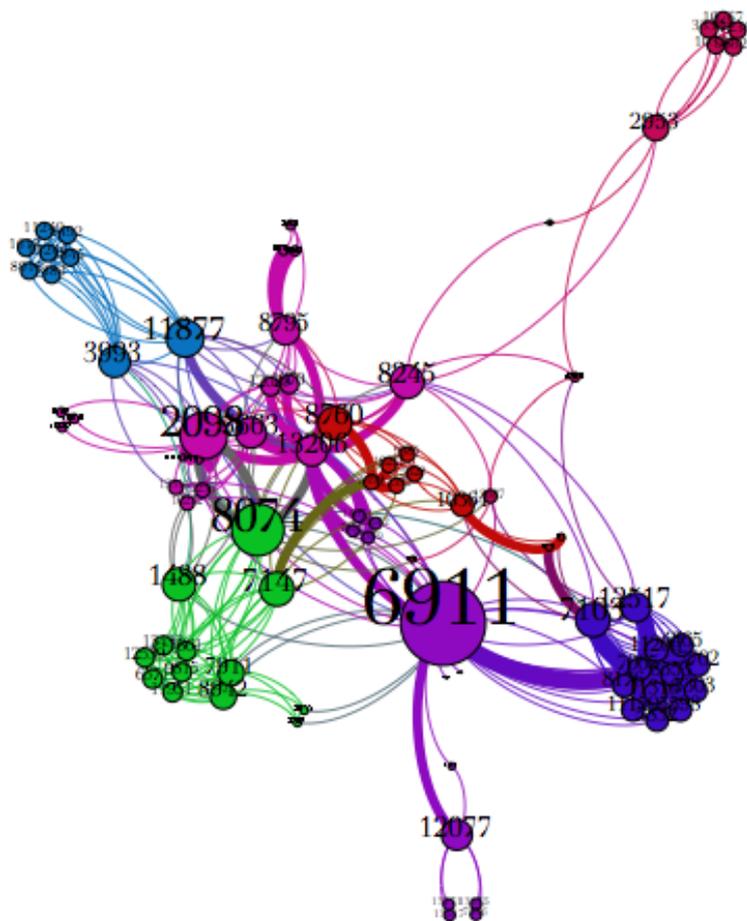


Figure 4. Graph of the filtered Al-Faloja users with posts related to 'Judaism'. Nodes sized according to PageRank.

‘Jihad’

My next search term was ‘Jihad’. By sampling I noticed from the context of the messages that usually the violent sense of the word is meant⁴⁵. On the Al-Faloja forum regarding this term I found intensive interaction in smaller

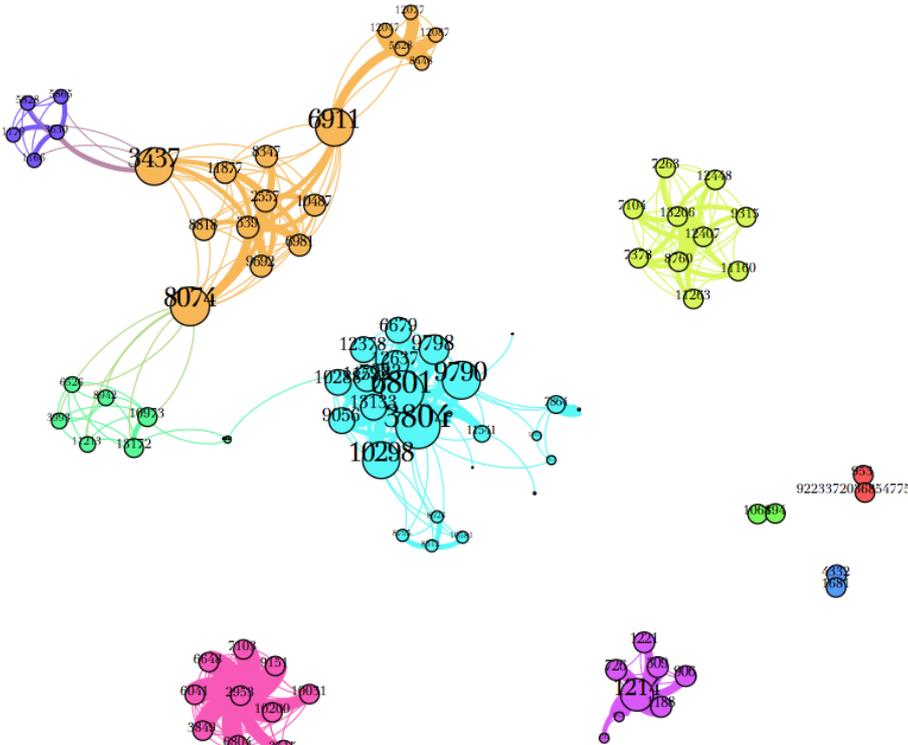


Figure 5. Graph of Al-Faloja forum users mentioning the word ‘Jihad’. Node sizes according to PageRank.

average degree is medium high when compared to other keywords, suggesting a moderate interest. Remarkable is the diameter of 8 steps, which is a lot for a network, especially for a network with people talking about the same subject. The users apparently are quite divided, which can also be seen from the graph: the different fragments are very thinly, if at all, connected to each other.

Within the components we see that in groups turquoise and especially pink there is quite some internal interaction, but in other groups this is not the case. An explanation could be that there is a different understanding of ‘jihad’: the subject might divide the users more than unite them. It is hard to tell what kind of brokerage is taking place in such a fragmented network. If anything, it should be coordinating brokerage, as can be seen with user #2953, central in group pink. In the other notable groups, turquoise and orange we see different patterns. The three larger orange nodes represent users with mostly coordinating roles, but the ties they facilitate are

(~8 members) and very much separated groups. In all groups a ‘leader’ was visible: one user with more posts than the other group members, but also with stronger ties to the other group members than they have amongst each other, for example the user in the middle of the pink or purple groups. The total picture is that of a strongly fragmented social network with very little or no interaction between the different components. The

⁴⁵ The word ‘Jihad’ has two meanings: the ‘greater’ and ‘lesser’ jihad. The former is an inner struggle for submission to Islam, the latter is the violent version against the enemies of the Islam.

somewhat thin. Group turquoise comes very close to the social idea of a clique, a dense group with much internal connectivity, and with the strength of ties equally distributed. Given the relatively high PageRank scores in group turquoise the composition of the group is very much susceptible to radicalization: high degrees and high internal density with few influence from outside.

On the Ansar forum the word is mentioned in 277 of the 11,244 threads, that is, 1 in every 50 threads, which is a high percentage for such a specific term (Bermingham 2008: 4). The average degree is 11.746, meaning that every user on average posted almost 12 messages regarding this subject, which is also very high. Apparently the topic has the interest of a certain group. Despite that, I have not really found cliques, although there are sorts of communities in this network (e.g. group green). Within this community there was too little interaction to call it a clique – only a few mutual threads were ‘jihad’ was discussed. The network is denser, but less clustered, and thus also less excluding. Within the whole network of 56 users there were only 4 that stood out as prominent users, both with more posts and more advantageous positions in the network. Most of the traffic was in the direction of the prominent users. The four prominents, being the larger nodes in red, blue and purple, however, all have strong ties among each other and so they actually form a clique themselves, although the modularity algorithm does not recognize them as such. They will very likely exchange ideas among one another.

Of the four prominents one user stands out, with a degree of 45 that is four times the average and having ties with 82 per cent of all nodes in the network, as clearly is evident from the graph. Such an extent of centrality is rare, and with a relatively small sized forum may suggest that the actor has a particular interest in jihad, or, indeed plays a coordinating role in recruiting people. I have sized the nodes according to betweenness, to show you how well users have been able to communicate with users from other clusters than their own, that is, how an exchange of ideas outside one cluster was possible. You can see that this is hardly the case, except for that one user. However, by filtering out the red user the average degree drops from 11.746 to 10.484 but still is quite

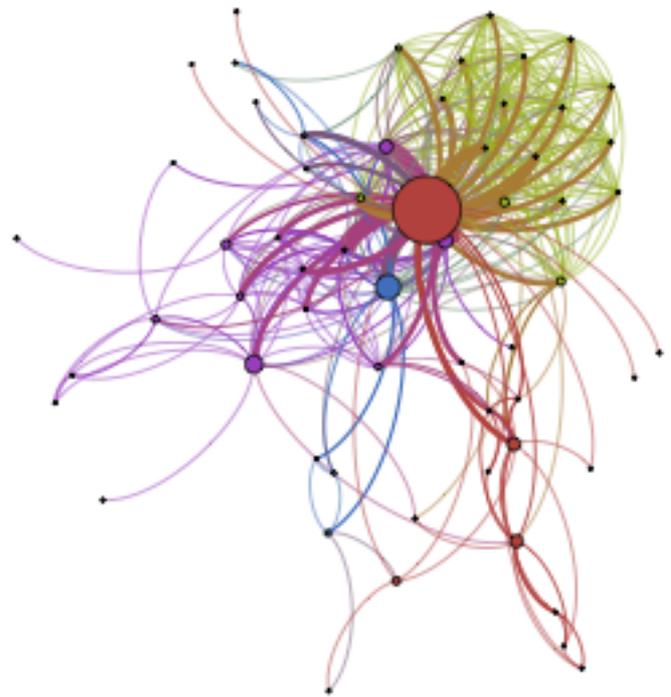


Figure 6. Graph of Ansar users mentioning the word 'jihad'. Nodes sized according to betweenness centrality.

high, showing an overall commitment to this topic. This situation is a liable climate for radicalization: loose, but open groupings and one very influential actor that is in contact with a vast majority of the users. The interaction between user red and the green group of users represents this pattern very clearly.

‘Mujahidin’

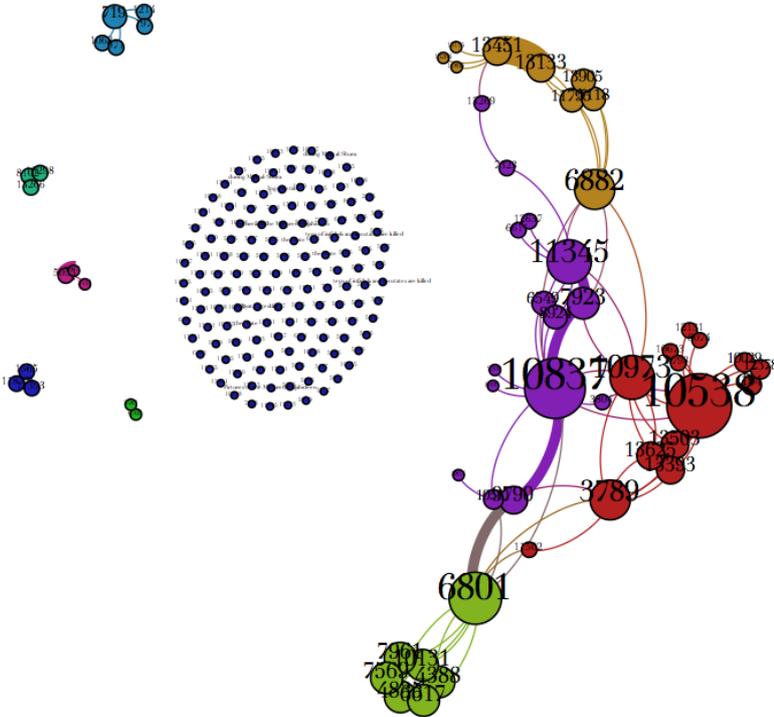


Figure 7. Graph of Al-Faloja users mentioning the word ‘Mujahidin’. The degree filter has been turned off.

to each other. The bulk of the messages were written in less than two months. Turning off the degree-filter showed that a considerable number of users have mentioned ‘mujahidin’ once, as you can see from the navy blue nodes in the graph. In fact, ‘mujahidin’ has the largest portion of one-message threads of all my samples⁴⁷. A low average degree and clustering coefficient further confirms the suggestion that with regard to ‘mujahidin’ no real (radicalizing) discussion is taking place, but that the word is not rarely mentioned. A sample content analysis showed that ‘mujahidin’ is mentioned mainly in news messages and video-related threads. Since May 2007, the word has been mentioned on 4,210 Al-Faloja threads.

The same pattern of many one-message threads I found on the English Ansar forum: more than 70 per cent of the messages were filtered out when

With ‘mujahidin’⁴⁶, on Al-Faloja both density and average degree are very low. Yet a number of actors (~3 to 7) have very high degree scores. However there are not very strong ties among the actors, suggesting that there is little interaction or discussion around this subject, but rather one-thread news messages please note the pattern of relatively thin lines in the graph. A time frame analysis revealed that all messages were written between September 2008 and October 2009, where the ‘heavyweights’ came in only from September 2009. At the same time, mutual threads connected two larger clusters

⁴⁶ Arabic for ‘strugglers’ or ‘people doing Jihad’. Nowadays the term mostly refers to (para-) military groups with radical Islamic convictions.

⁴⁷ See also the Appendix 2.

selecting nodes with degree larger than 1. I noticed that for every search term the density on the smaller Ansar forum was always higher than on the larger Al-Faloja forum. There are no noteworthy cliques. In the ‘mujahidin’-graph the top users in terms of degree and PageRank are #10 and #12. They are also respectively the number 3 and 1 on the whole forum with the most posts. Another user that stands out is #1694. He has a considerable lower degree than the top users, but has many mutual threads with a small numbers of users on which the word ‘mujahidin’ is mentioned. He calls himself ‘mujahed88’ on the forum and appears to speak Arabic. The demeanour of his postings is authoritative and he has adequate knowledge of the Quran⁴⁸. Based on his posts I read and his position in the network, it is possible that he is one of those senior influential figures asserting influence over cliques, which Sageman mentioned (2008: 79).

When we look at the graph we see that 1694’s (in gold) position indeed is remarkable, in a typical coordinating brokerage role. He is, however, not the only one in such a position. User #932, just above him, in fuchsia, has extremely strong ties with only 6 users. There are a handful more users in such a position, but their messages are not directed toward a close-knit group, but rather toward a few individuals. There is no regularity in the distribution of strong ties. In terms of radicalization, this means that the social pattern of strong ties seems one of bi- or trilateral relations, but not of cliques.

Clustering coefficient and density are low. If radicalization is taking place, and the high occurrence of the word on the total forum (~5,16 per cent) point that way, it is not taking place through discussion but rather by videos and images, as evidenced by the general content of the messages.

General observations

What stands out is that actors with a high degree have more connections with each other than with users with a lower degree. It seems that there is

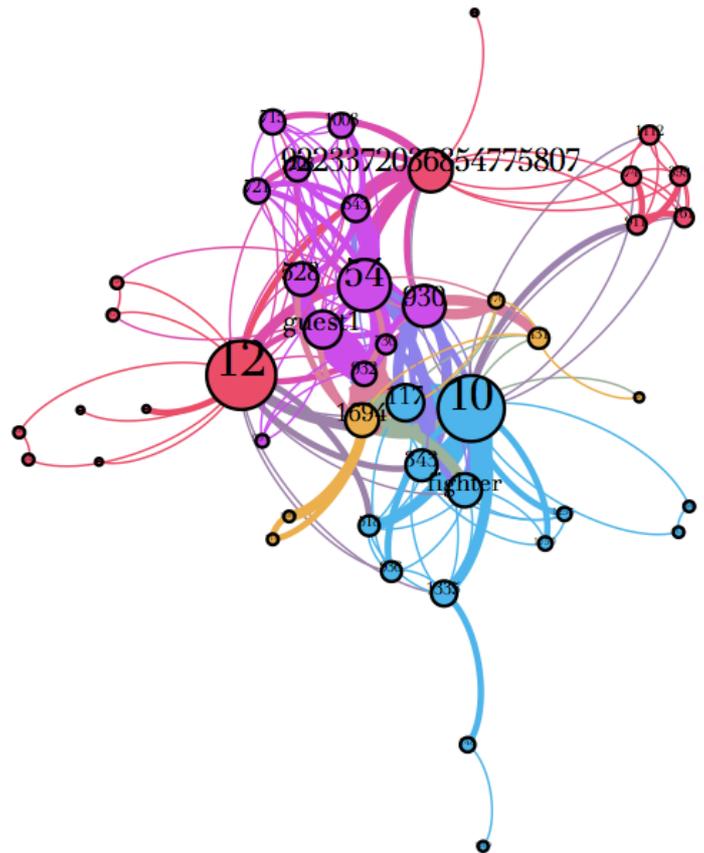


Figure 8. *As-Ansar forum users mentioning the word ‘mujahidin’*. Nodes sized according to PageRank.

⁴⁸ See DWFP Ansar1 thread #27839, #27750, #27173, #26070.

an elite community within the forum where there is more interaction within, than there is to the outside. It is the same prominent actors that figure as hubs between clusters, and that also have high betweenness centrality scores, meaning that they are in advantageous positions in the network. They have posted a lot and they have read a lot. I have experimented with different definitions of power and centrality to identify those actors.

I have been able to identify some community structures, including cliques, though not with every search term. In fact, cliques in the sense of tight social units occurred only a couple of times. What I found much more often was a group in which one user had stronger ties to other group members than the others had among themselves. A user like that typically did not have a higher degree than the others, pointing to an equivalent investment of time in the forum. Such an actor apparently is a much-wanted contact who enjoys the respect of his environment. As such he is in an advantageous position, and apt to brokerage. However, within a clique a sense of equality should be present, and therefore such groups I think cannot be considered as such.

This general observation is related to the patterns of brokerage. In networks that are very much decentralized and fragmented, brokerage roles tend to be that of a coordinator rather than a consultant. Brokerage also depends on the density, average degree and diameter of a network. These give an idea of how and how fast information is not travelling around. In well-connected parts of the network the role of the coordinator has bigger potential, whereas in less dense areas the consultant can assert more influence.

Lastly, I have investigated social structures on two different radical Islamic forums: Al-Faloja and As-Ansar, with Al-Faloja being fifteen times bigger. I found that the percentage of one-message threads was quite similar on both forums. As-Ansar showed more density, but less clustering, although this varies with different terms. Al-Faloja showed more signs of closer interaction, with more strong ties within clusters, whereas As-Ansar was more centralized – as could be expected with a smaller forum – but with weaker ties. I have given some examples of the discourse on both forums related to the keywords. It seems both forums have a potential for radicalization, but in different ways: through discussion, links to videos, images, newsflashes, to name but a few.

Serendipity: side-findings worth mentioning

I have been able to experiment with a timeline module, a feature of network analysis that has become publicly available only a few months ago⁴⁹. It allows

⁴⁹ Thanks, once more, to Sascha Ramondt who added this feature, that is not yet available on the DWFP, to the dataset.

me to see at what time users became active on the forum and were adopted in the network. I have used this for some purposes of my research, but I have not used the full potential of it: the development of user posts and threads over a period of time, which may lead to finding interesting new correlations with this new dimension added to the network. For example, it is known that the Al-Faloja forum is, among other things, concerned with the war and insurgency in Iraq. In the timeline module I see a sharp rise in the mentioning of the word ‘crusader’ in August and September of 2009. It peaked when the number of messages on the whole forum was in decline. It seems that the word was used significantly more often in that period. In the late summer of 2009, after control over Baghdad was ceded to the Iraqi government, two extremely bloody attacks occurred, one on August 19th and one on October 25th, with the attack in August being the deadliest of the year. It is likely that there is a correlation between the increase of radical forum activity and the attacks. It is impossible on a forum with more than 1000 posts a day to literally check and read what users have discussed, but it is likely that there is a connection. In future research it could be very interesting to see what the influence of developments in the Iraq war on forum activity is, or perhaps, even more interesting, vice versa. It is, to the best of my knowledge, not yet possible to view the historical course of prominence of the nodes, but this could also be an interesting addition⁵⁰.

⁵⁰ See also my future recommendations at the end of the conclusion.

7. Conclusion

The goal of my research was to identify social patterns of radicalization within two of the 28 Islamic forums worldwide that have been identified as radical, extremist, and possibly terrorist⁵¹. Although there is no collective agreement, many scholars have acknowledged that Internet forums are a place where radicalization occurs, and thus form a pool of recruitment for terrorist organizations (Bermingham et al. 2008; Weimann 2008). Sageman, an important and one of the only social scholars on the field of radicalization, has argued that radicalization, usually happens clique-wise. That is, radicalization is a group process, rather than an individual process. Sageman, however, did not mean online radicalization on Internet forums but pointed instead to fundamentalist mosques as being nests of radical Islamic thought (2008: 143). From another corner of science, Security Informatics, there is also interest in the process of radicalization. That interest is mainly aimed at the identification of online centres of radicalization. Identifying such centres can best be achieved using automated web mining approaches: an advanced tool that searches, identifies, stores and labels online radicalizing content, especially forums. This has resulted in an enormous treasure of information. The Artificial Intelligence Lab of the University of Arizona plays a leading role in this young branch of information science and runs the Dark Web Forum Portal. I have been granted permission to use this portal for my research. My methodology was aimed at synthesizing these two approaches – the social approach with Sageman and the informatics approach. Inspired by Hanneman’s handbook (2005) I have designed a method to first identify cliques; second, identify prominent figures; and finally to search for interaction between those two social entities. I have written about my findings and interpreted them. In this final chapter I will evaluate my findings and interpretations from the previous chapter and discuss whether or not I have found what I was looking for. Further, I will discuss the meaning of my approach for the field of Conflict Studies as well for Security Informatics. I will conclude with some future recommendations for further research.

Using different definitions – most notably degree, betweenness centrality and PageRank – I have been able to identify prominent figures. I have argued why in the situation under scope some actors are more influential than others. I have also argued why different definitions of being influential apply to different situations. Using a partition algorithm called ‘modularity’ the graph was split into groups. Alas, I found that the basis for representing those groups as cliques in most cases was quite thin. In strict terms cliques were present: there were enough groups in which all members have ties to

⁵¹ DWFP, Home: List of Forums. See also Appendix 1.

every other member. However, there was not enough interaction within the group to be able to speak of a ‘clique’ in a way that socially makes sense indeed, one or two mutual threads is not a realistic definition of a clique. And so I think I did not find enough evidence to convincingly state that my earlier statement has been exactly proven. An adapted version of it, however, I think is.

One reason for this is the presence of a lot of noise. Whatever search terms I used, always more than half of the search results were one-message threads. Although this also says something, for the purpose of my research those messages are not useful because there is no proof of interaction. Therefore the size of my samples was reduced, and then it is also harder to find patterns. Finally, my research has an exploratory character. I cannot definitely prove that radicalization is taking place, which is extremely hard. I am however able to prove that the social climate makes is highly plausible. Even if I did find convincing evidence of a social climate prone to clique-wise radicalization, other methods still would be necessary to verify that it has actually taken place.

I did find a number of interesting social patterns. But first it is important to recognize that the pre-selection of forums by the Dark Web Forum Portal already is evidence of the fact that radicalization is definitely happening on those forums. The solid methodology of assembling the collection⁵² assures that. Also, it is legitimate to assume that every member of a radical Islamic forum has an interest in violent jihad, be it actively, passively, verbal or violent; otherwise they would not have signed up to become a member of that forum.

I found that instead of an entire clique only one clique member has a lot of interaction with a powerful actor, as evidenced by a large number of mutual threads. This pattern was present with several keywords on both the Al-Faloja and Ansar forums. Within the clique, however, this person often had more ties to other clique members than the other clique members among themselves. There was no equal internal interaction in the clique. This is not the social idea of a clique: equivalency should be the basis. Within a clique, of course there can be a leader or a *primus inter pares*, but not to the extent which I found. So what I call cliques here actually are more cores or components around that one user with a significant higher degree than the rest. Because the (online) interaction within is limited because of that I cannot accept my earlier hypothesis.

Another thing I found was the high amount of interaction among prominent forum users. Generally, I think that radicalization is *most* occurring under those prominent forum users – those with high degree, PageRank, and

⁵² See also chapter 4.

betweenness centrality scores. Those users are very active on the forum: the top users on the Al-Faloja forum posted on average more than 1,100 messages per month⁵³. This indicates that they have invested a considerable amount of time in being active on radical Islamic forums. Again, the fact that someone is very active on an extremist Islamic forum already is an important indicator for radicalization with that person. So if someone is extremely active this proves his commitment. There are no signs of a certain hierarchy among them, so it could be that the interaction between them is a cycle of self-radicalization, confirmation, and encouragement, just like with Mohammed B. Furthermore, I have pointed at a number of statistics that give a lot of information about how fast information can travel through a network; and how different combinations of strong or weak density, diameter, and average degree affect the potential for different types of influencing, most notably by consulting and coordinating. Centralization and fragmentation are also important aspects with regard to radicalization: more fragmented networks tend to encompass more cliques, which according to other social research remain the most important social unit to observe (Sageman 2008: 154).

I think my modest effort to try to synthesize Informatics and Conflict Studies was not unfruitful. The concept of looking at processes of radicalization, which these days are important root causes of conflict, in a relational way, and to overlook the entire process of interaction can be useful for our field. Because the approach is in principle quantitative, it encompasses the whole, and so you can very easily get a good picture of the overall process of interaction, identify the heavyweights, and see where the centres of interaction are. Those are of key importance when studying social group processes in whatever context. Plus, within the domain of addressing Internet interaction as either cause or catalyst of radicalization, (semi-) automated data collection techniques are essential, and social network analysis a great tool if we are to make sense of radicalization on the Internet, the actuality of which sadly is evermore proved by examples such as Mohammed B., Anders Breivik, Al-Qaeda, etcetera. From the Security Informatics point of view then, I think incorporating insights on group interaction from social sciences can be very helpful in gaining a better understanding, and has yet, to the best my knowledge, been done too little. By synthesizing the both approaches we can arrive at the meso-level of analysis: not the individual account, neither the whole picture, but the interaction between and within groups.

Expanding time frame analysis could enhance future research in this young field. This adds an extra dimension to the social network analysis, which enables the researcher to evaluate the development of the network, and see what events caused forum activity and interaction to increase or decrease.

⁵³ DWFP, Al-Faloja member list

This could be extremely interesting because it may predict circumstances under which radicalization is occurring heavier than usual, and thus help it from escalating. But also socially it can be interesting: it may say something about the process of preferential attachment by which networks are expanded⁵⁴, and give us a better understanding of this dynamic group process. A fragile part of the whole method remains the searching by keyword as most important selection criterion, especially in this specific lexical domain. Limited knowledge of Arabic languages may cause researchers to overlook important themes, as well as not understanding metaphors or other expressions for which more advanced knowledge of a language is required. Therefore automated data collection techniques should always be developed in close cooperation with linguistics, to ensure that lack of language skills affect the results to the minimum.

Despite that no agreement among academic scholars has been reached, I think online radicalization in the 21st century is a fact. The influence of Internet in our daily lives is enormous. With examples like the ones I just gave, at the very least it is impossible to uphold the claim that radicalization and mobilization are *not* happening on the Internet. When focusing the attention on forums as being only one of a multitude of ways in which terrorist organizations are present on the net (think of web sites, chat rooms, videos, YouTube) identifying central actors is not so difficult as they distinguish themselves easily, usually by a high number of posts. Having those actors identified enables intervention and guidance easier. This can prevent those people from radicalizing further, if there is a will to do so. The technique, however, is not all-embracing, and remains a cooperation between man and machine, between data collection and interpretation.

I hope to have contributed to a better understanding of online radicalization. I think I have introduced a largely unexplored area and some new analytic tools to the field of Conflict Studies. An area, also, that will very likely increase in importance in the near future. With an improved understanding, perhaps violent excesses could be prevented.

I thank you for reading my thesis.

Lucas Lenselink

⁵⁴ See chapter 3.

Appendix 1

List of the 28 forums monitored within the Dark Web Forum Portal

Name	Messages	Threads	Users	Period active ⁵⁵	Description ⁵⁶ (DWFP)
Al-Boraq www.alboraq.info	270,048	65,715	3,858	Jan 2006 Dec 2010	This Arabic forum was established as an affiliation of the Jihad and Reform Front, an umbrella organization consisting of the Islamic Army in Iraq and other groups.
Al-Fallujah Islamic Forums [Al-Faloja] (DEAD)	564,159	76,940	5,881	Sep 2006 Jan 2010	This Arabic forum, which includes an English-language section, is one of the most popular and credible among jihadists. It supports the insurgency in Iraq as well as the global jihad.
Al-Firdaws Forum (DEAD)	39,715	9,359	2,187	Jan 2005 Dec 2007	A general Arabic forum with subsections containing discussions of Islamic ideologies and generally supporting Jihadi organizations.
Midad Al-Suyuf (DEAD)	40,248	11,587	1,685	Mar 2006 Sep 2010	This Arabic forum hosts statements from Al-Qaeda and affiliated groups, Iraqi insurgents, and jihadists worldwide. It includes an English-language section.
Alokab www.alokab.com/forums	82,231	12,419	1,699	Apr 2005 Dec 2010	An Arabic forum dedicated to Islamic theology; contains some radical content.
Al-Qimmah Islamic Network www.alqimmah.net	31,793	16,199	367	Nov 2007 Dec 2010	This Sweden-based forum disseminates videos and statements from jihadist groups worldwide, primarily Somalia, Iraq and Afghanistan. Includes a mixed-language section on computers and the internet.
Alsayra www.alsayra.com/vb	1,283,393	155,935	72,866	Apr 2001 Dec 2010	An Arabic forum associated with the Sayra online newspaper; some members of the forum sympathize with the causes of more extremist Islamist groups.
Ansar Al-Jihad Network www.as-ansar.com/vb	50,803	12,949	1,326	Nov 2008 Feb 2010	This Arabic Jihadist forum has English and German sites that are popular with Western jihadists and provide invitation-only sections.
At-Tahadi www.atahadi.com/vb	5,484	2,619	322	Apr 2008 Dec 2010	This Arabic forum covers jihadist insurgent groups in Iraq and al-Qa'ida-affiliated groups elsewhere, as well as the Taliban.
Hanin Net www.hanein.info/vb/main	988,032	125,211	3,177	Nov 2006 Nov 2010	Dedicated to news on jihad and Iraqi insurgent groups, this Arabic forum

⁵⁵ Generally forum activity until November or December 2010 has been measured. Some forums are still active. If so the hyperlink is added under the forum name.

⁵⁶ The descriptions are all cited from the Dark Web Forum Portal homepage: http://128.196.40.222:8080/CRI_Indexed_new/index.jsp

.php						includes an English-language section. It supports both Iraqi nationalist and jihadist groups operating in Iraq.
Hawaa World forum.hawaaworld.com/index.php	2,458,126	43,341	118,776	Jan 2001 May 2010		An Arabic forum dedicated to Muslim women; some members sympathize with and support radical Islamists.
Hadramout www.hdrmut.net/vb	1,552,227	151,694	29,491	Nov 2000 Dec 2009		A Yemeni Arabic forum with registered members who sympathize with the causes of extremist Islamist groups.
Ma'arik Islamic Network (DEAD)	60,417	16,685	1,903	Jul 2007 Oct 2010		This forum covers the insurgency in Iraq and global jihad. It includes an English-language section and technical sections that cover web development, encryption, computers, and hacking.
Al Mujahidin Electronic Network (DEAD)	165,510	36,010	4,867	Nov 2007 Nov 2010		This Arabic forum covers the global jihad. It features an English section and a technical section on computers and the Internet, which has included postings about hacking and calls for cyber attacks.
Montada www.montada.com	1,478,167	127,903	43,530	Sep 2000 Jul 2010		A general Arabic forum from the Gulf region. Discussions cover a wide range of diverse topics such as supporting Islamist militants and computer software.
Ana al-Muslim www.muslm.net/vb	1,603,518	204,602	12,722	Oct 1996 Nov 2010		Created in 1996, this popular forum is one of the oldest Salafi forums. It is more Salafi than jihadi and covers a wide variety of viewpoints, but it has been used by al-Qaeda.
Shumukh al-Islam Network (DEAD)	299,956	48,472	4,059	Mar 2007 Feb 2010		Established in April 2007, this popular Arabic forum includes an English section and distributes news on Iraqi insurgent groups such as the Islamic State of Iraq and global Al-Qa'ida affiliates.
Ansar al-Jihad Network (English) (DEAD)	29,492	11,244	382	Dec 2008 Jan 2010		The English site of Ansar al-Jihad Network that is popular with Western jihadists; it provides invitation-only sections.
Gawaher www.gawaher.com	579,849	212,772	8,026	Oct 2004 Nov 2010		An English language Islamic forum dedicated to discussions of issues pertaining to the Islamic world and Islam. Some forum members sympathize with radical Islamic groups.
Islamic Awakening forums.islamicawakening.com	129,425	27,968	2,803	Apr 2004 May 2010		An English language Islamic forum with members based in the UK and other countries.
Islamic Network talk.islamienetwork.com	91,874	13,995	2,082	Jun 2004 Nov 2010		An English-language forum dedicated to various topics of interest to Muslims, from theology to discussions of recent world events. Some registered members sympathize with and support terrorist organizations.

Islamic Web Community www.myiwc.com/forums/index.php	25,016	6,310	756	Nov 2000 Feb 2010	An English language forum serving Muslim communities. Most English language forums, including this one, are targeted to non-Arabic speaking Muslims as evidenced by threads dedicated to learning the Arabic language.
Turn To Islam www.turntoislam.com/forum	327,542	40,825	10,527	Jun 2006 Nov 2010	As announced on the homepage of the forum, this is an English language forum with the goal of "correcting the common misconceptions about Islam". Radical participants may occasionally display their support for fundamentalist militant groups.
Ummah www.ummah.com/forum	1,357,525	82,674	16,941	Apr 2002 Nov 2010	An English language Islamic forum serving Muslim communities. As is the case with other Islamic forums, full respect and abiding by Islamic law is expected from all forum participants.
Al-Minha Dj forum-alminhadj.com	7,849	2,643	392	Jun 2008 Dec 2010	A French language Islamic forum dedicated to the acquisition of knowledge of religion, through shared resources and participation.
Forums d'Aslama www.aslama.com/forums	143,576	22,236	2,967	Oct 2004 Jul 2010	A French language Islamic forum for sharing, teaching and living the religion of Islam.
Al-Mourabitoune (DEAD)	72,140	7,905	3,198	May 2002 Mar 2008	A French language Islamic forum for sharing life experiences and ideologies among Islamic groups.
Ansar al-Dschihad Netzwerk (German) (DEAD)	1,768	766	71	Feb 2009 Jan 2010	The German site of Ansar al-Jihad Network that is popular with Western jihadists and provides invitation-only sections.
KavkazChat www.kavkazchat.com	558,042	6,144	5,634	Mar 2003 Jan 2010	This Russian-language forum focuses on jihad in the North Caucasus. It includes a cyberwarfare section that supports the Ansar cyber unit with cyber attacks.

Appendix 2
Basic Graph Statistics

Forum	Keyword	Nodes ⁵⁷	Edges	Density ⁵⁸	Diameter	Average Degree	Average Clustering Coefficient
Al-Faloja	Crusader	108 (54.82%)	841	0.146	4	15.547	0.879
	Jihad	88 (44.00%)	328	0.086	8	7.455	0.799
	Mujahidin	61 (30.50%)	116	0.063	6	3.803	0.664
	Judaism	107 (53.50%)	406	0.072	6	7.589	0.796
As-Ansar [EN]	Crusader	20 (46.51%)	47	0.247	4	4.700	0.730
	Jihad	63 (48.84%)	363	0.236	4	11.746	0.748
	Mujahidin	46 (31.29%)	180	0.174	5	7.826	0.723

⁵⁷ Absolute and percentage of total nodes. Nodes with a degree less than 1 have been left out. Naturally, all possible edges are included.

⁵⁸ Graph density, diameter, average degree, and average clustering coefficient after the same filter has been applied.

Bibliography

Primary sources

ONLINE ARCHIVAL SOURCE

Artificial Intelligence Lab, University of Arizona. (2008-2010). *Dark Web Forum Portal*. Retrieved August 2011, from Dark Web Forum Portal: http://128.196.40.222:8080/CRI_Indexed_new/index.jsp

SOFTWARE

Gephi Contributors. (2008-2011). Gephi Graph Visualization and Manipulation Software. *Gephi 0.8 alpha* .

Secondary sources

PUBLISHED SOURCES

Benschop, A. (2005). *Jihad in Nederland. Kroniek van een aangekondigde politieke moord*. (Universiteit van Amsterdam) Retrieved July 2011, from Eigenaardigheden van Cyberspace: <http://www.sociosite.org>

Bermingham, A., Conway, M., McInerny, L., O'Hare, N., & Smeaton, A. F. (2008). *Combining Social Network Analysis and Sentiment Analysis to Explore the Potential for Online Radicalisation*. CLARITY Centre for Sensor Web Technologies and School of Law and Government, Dublin City University. Dublin: School of Law and Government, Dublin City University.

Chen, H., Chung, W., Qin, J., Reid, E., Sageman, M., & Weimann, G. (2008). Uncovering the Dark Web: A Case Study of Jihad on the Web. *Journal of the American Society for Information Science and Technology* , 59 (8), 1347-1359.

Chen, H., Reid, E., Sinai, J., Silke, A., & Ganor, B. (Eds.). (2008). *Terrorism Informatics*. New York: Springer Science & Business Media.

Conway, M. (2005). Terrorist 'Use' of the Internet and Fighting Back. *Cybersafety: Safety and Security in a Networked World: Balancing Cyber-Rights and Responsibilities* (pp. 1-34). Oxford: Oxford University.

Hanneman, R. A., & Riddle, M. (2005). *Introduction to Social Network Methods*. Riverside, California: University of California.

- Hoffman, B. (2008). The Myth of Grass-Roots Terrorism. *Foreign Affairs* , 87.
- Katz, R., & Devon, J. (2007). A Web of Terror. *Forbes* , 179 (10), 184-188.
- Qin, J., Zhou, Y., Reid, E., Lai, G., & Chen, H. (2007). Analyzing Terror Campaigns on the Internet: Technical Sophistication, Content Richness, and Web Interactivity. *International Journal of Human-Computer Studies* , 65, 71-84.
- Qin, J., Xu, J. J., Hu, D., Sageman, M., & Chen, H. (2005). Analyzing Terrorist Networks: A Case Study of the Global Salafi Jihad Network. *ISI 2005, LNCS 3975* (pp. 287-304). Berlin: Springer Verlag.
- Reid, E., & Chen, H. (2007). Mapping the contemporary terrorism research domain. *International Journal of Human Computer Studies* , 65, 42-56.
- Reid, E., Qin, J., Zhou, Y., Lai, G., Sageman, M., Weimann, G., et al. (2005). *Collecting and Analyzing the Presence of Terrorists on the Web: A Case Study of Jihad Websites*. Tucson: University of Arizona.
- Rodriguez, J. A. (2005). *The March 11th Terrorist Network: In Its Weakness Lies Its Strength*. Grupo de Estudios de Poder y Privilegio, Departament de Sociologia i Anàlisi de les Organitzacions. Barcelona: Universitat de Barcelona.
- Saddiq, M. A. (2010). *Whither e-Jihad: Evaluating the Threat of Internet Radicalization*. S. Rajaratnam School of International Studies. Singapore: Nanyang Technological University Library.
- Sageman, M. (2004). *Understanding Terror Networks*. Philadelphia, Pennsylvania: Pennsylvania University Press.
- Sageman, M. (2008). *Leaderless Jihad. Terror Networks in the 21st Century*. Philadelphia, Pennsylvania: University of Pennsylvania Press.
- Salem, A., Reid, E., & Chen, H. (2006). Content Analysis of Jihadi Extremist Groups' Videos. In S. Mehotra (Ed.), *ISI 2006, LNCS 3975* (pp. 615-620). Berlin: Springer Verlag.
- Stohl, M. (2006). Cyber terrorism: a clear and present danger, the sum of all fears, breaking point or patriot games? *Crime Law Social Change* , 46, 223-238.
- Thomas, T. L. (2003). Al Qaeda and the Internet: The Danger of "Cyberplanning". *Parameters* , 112-123.

- Weimann, G. (2004). *www.terror.net. How Modern Terrorism Uses the Internet*. Washington D.C.: United States Institute of Peace.
- Weimann, G. (2008). The Psychology of Mass-Mediated Terrorism. *American Behavioral Scientist*, 52 (69), 69-86.
- Xu, J., Chen, H., Zhou, Y., & Qin, J. (2006). On the Topology of the Dark Web of Terrorist Groups. In S. Mehotra (Ed.), *ISI 2006, LNCS 3975* (pp. 367-376). Berlin: Springer Verlag.
- Zhou, Y., Qin, J., Lai, G., Reid, E., & Chen, H. (2006). Exploring the Dark Side of the Web: Collection and Analysis of U.S. Extremist Online Forums. In S. Mehotra (Ed.), *ISI 2006, LNCS 3975* (pp. 621-626). Berlin: Springer Verlag.