

Analyzing terror campaigns on the internet: Technical sophistication, content richness, and Web interactivity

Jialun Qin^{a,*}, Yilu Zhou^b, Edna Reid^c, Guanpi Lai^d, Hsinchun Chen^c

^aDepartment of Management, University of Massachusetts Lowell, Lowell, MA 01854, USA

^bInformation Systems and Technology Management, George Washington University, Washington, DC 20052, USA

^cDepartment of Management Information Systems, The University of Arizona, Tucson, AZ 85721, USA

^dSystems and Industrial Engineering Department, The University of Arizona, Tucson, AZ 85721, USA

Available online 1 November 2006

Abstract

Terrorists and extremists are increasingly utilizing Internet technology to enhance their ability to influence the outside world. Due to the lack of multi-lingual and multimedia terrorist/extremist collections and advanced analytical methodologies, our empirical understanding of their Internet usage is still very limited. To address this research gap, we explore an integrated approach for identifying and collecting terrorist/extremist Web contents. We also propose a Dark Web Attribute System (DWAS) to enable quantitative Dark Web content analysis from three perspectives: technical sophistication, content richness, and Web interactivity. Using the proposed methodology, we identified and examined the Internet usage of major Middle Eastern terrorist/extremist groups. More than 200,000 multimedia Web documents were collected from 86 Middle Eastern multi-lingual terrorist/extremist Web sites. In our comparison of terrorist/extremist Web sites to US government Web sites, we found that terrorists/extremist groups exhibited similar levels of Web knowledge as US government agencies. Moreover, terrorists/extremists had a strong emphasis on multimedia usage and their Web sites employed significantly more sophisticated multimedia technologies than government Web sites. We also found that the terrorists/extremist groups are as effective as the US government agencies in terms of supporting communications and interaction using Web technologies. Advanced Internet-based communication tools such as online forums and chat rooms are used much more frequently in terrorist/extremist Web sites than government Web sites. Based on our case study results, we believe that the DWAS is an effective tool to analyse the technical sophistication of terrorist/extremist groups' Internet usage and could contribute to an evidence-based understanding of the applications of Web technologies in the global terrorism phenomena.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Web content analysis; Web usage analysis; Web collection building

1. Introduction

The weekly news coverage of excerpts from messages and videos produced and Web-cast by terrorists/extremists has shown that terrorists and extremists have become exploiters of the Internet beyond routine communication operations. Internet has dramatically increased their ability to influence the outside world. Several virtues of the

Internet, such as ease of access, anonymity of posting, huge audience, and lack of regulations, have enabled terrorists to directly speak to millions of people—both supporters and adversaries, with little chance of being detected. As posited by Jenkins (2004), through operating their own Web sites and online forums, terrorists have effectively created their own “terrorist news network.”

Terrorist/extremist organizations have generated thousands of Web sites that support psychological warfare, fundraising, recruitment, coordination, and distribution of propaganda materials. From those terrorist/extremist Web sites, supporters can download multimedia training materials, buy games, T-shirts, and music CDs, and access forums and chat services such as PalTalk (Bowers, 2004;

*Corresponding author. Fax: +1 978 9343011.

E-mail addresses: qin@email.arizona.edu, jialun.qin@gmail.com (J. Qin), yilu@email.arizona.edu (Y. Zhou), edanread@eller.arizona.edu (E. Reid), guanpi@email.arizona.edu (G. Lai), hchen@email.arizona.edu (H. Chen).

Muriel, 2004; Weimann, 2004). Some Web sites such as those associated with the Jihad terrorist/extremist movement are extremely dynamic in that they emerge overnight, frequently modify their contents, and then swiftly ‘disappear’ by changing their URLs which are later announced via online forums (Weimann, 2004). They are often hosted on free Web space servers or by unsecured and poorly maintained commercial servers. Such Web sites are technically supported by those who are Internet Savvy to provide sophisticated propaganda images and videos via proxy servers to mask ownerships (Armstrong and Forde, 2003). The level of technical sophistication of the Islamic terrorist/extremist organizations’ Web sites has increased according to Katz, who monitors Islamic fundamentalist Internet activities (Internet Haganah, 2005). The rapid proliferation and increased sophistication of Web sites and online forums run by terrorist/extremist organizations are indications of the growing popularity of the Internet in terrorism campaigns. They also indicate that there is a vast pool of sympathizers that such organizations have attracted, with some applying their IT expertise as contributions to the cause (Jesdanun, 2004).

Although this alternate side of the Internet, referred to as the “Dark Web,” has received extensive government and media attention, there is a dearth of empirical studies that examine the sophistication of terrorist/extremist organizations’ Web sites and how they support strategic and tactical information operations. Therefore, some basic questions about terrorist/extremist organizations’ Internet usages remain unanswered. For example, what are the major Internet technologies that they have used on their Web sites? How sophisticated and effective are the technologies in terms of supporting communications and propaganda activities?

In this study, we explore an integrated approach for collecting and monitoring terrorist-created Web contents and propose a systematic content analysis approach to enable quantitative assessment of the technical sophistication of terrorist/extremist organizations’ Internet usages. The rest of this paper is organized as follows. In Section 2, we briefly review previous works on terrorists’ use of the Internet. In Section 3, we present our research questions and the proposed methodologies to study those questions. In Section 4, we describe the findings obtained from a case study of the analysis of technical sophistication, content richness, and Web interactivity features of major Middle Eastern terrorist/extremist organizations’ Web sites and a benchmark comparison of Middle Eastern terrorist/extremist Web sites and Web sites from the US government. In the last section, we provide conclusions and discuss the future directions of this research.

2. Literature review

2.1. Terrorism and the internet

Previous research showed that terrorists/extremists mainly utilize the Internet to enhance their information

operations surrounding propaganda, communication, and psychological warfare (Thomas, 2003; Denning, 2004; Weimann, 2004). To achieve their goals, terrorists/extremists often need to maintain a certain level of publicity for their causes and activities to attract more supporters. Prior to the Internet era, terrorists/extremists maintained publicity mainly by catching the attention of traditional media such as television, radio, or print media. This was difficult for them because terrorists/extremists often could not meet the editorial selection criteria of those public media (Weimann, 2004). With the Internet, terrorists/extremists can bypass the requirements by traditional media and directly reach hundreds of millions of people, globally—24/7.

Terrorist/extremist groups have sought to replicate or supplement the communication, fundraising, propaganda, recruitment, and training functions on the Internet by building Web sites with massive and dynamic online libraries of speeches, training manuals, and multimedia resources that are hyperlinked to other sites that share similar beliefs (Coll and Glasser, 2005; Weimann, 2004). The Web sites are designed to communicate with diverse global audiences of members, sympathizers, media, enemies, and the public (Weimann, 2004). Table 1 summarizes terrorist/extremist groups’ objectives and tasks that are supported by Web sites.

2.2. Existing Dark Web studies

In recent years, there have been studies of how terrorists/extremists use the Web to facilitate their activities (Zhou et al., 2005; Chen et al., 2004; ISTS, 2004; Thomas, 2003; Tsfati and Weimann, 2002; Weimann, 2004). For example, researchers at the Institute for Security Technology Studies (ISTS) have analysed dozens of terrorist/extremist organizations’ Web sites and identified five categories of terrorists’ use of the Web: propaganda, recruitment and training, fundraising, communications, and targeting. These usage categories are supported by other studies such as those by Thomas (2003), Katz at SITE Institute (2004), and Weimann (2004).

Since the late 1990s, several organizations, such as SITE Institute, the Anti-Terrorism Coalition, and the Middle East Media Research Institute (MEMRI), started to monitor contents from selected terrorist/extremist Web sites for research and intelligence purposes. Tsfati and Weimann (2002) studied the content types and target audiences of terrorist/extremist organizations’ Web sites by analyzing the content of 29 Middle Eastern Web sites. Table 2 lists some of the organizations that capture and analyse terrorists/extremists’ Web sites grouped into three functional categories: archive, research center, and vigilante community.

Except for the Artificial Intelligence (AI) Lab, none of the enumerated organizations seem to use automated methodologies for both collection building and analysis of the Web sites. Due to the enormous size and the dynamic nature of the Web, the manual collection and analysis approaches have

Table 1
How web sites support objectives of terrorist/extremist groups

Terrorist/extremists' objectives	Tasks supported by web sites	Web features (Preece, 2000)
Enhance communication (Becker, 2004; Weimann, 2004)	<ul style="list-style-type: none"> ● Composing, sending & receiving messages; ● Searching for messages, information & people ● One-to-one, one-to-many communications ● Maintaining anonymity 	<ul style="list-style-type: none"> ● Synchronous (chat, video conferencing, MUDs, MOOs) & asynchronous (email, bulletin board, forum, UseNet newsgroup) ● GUI ● Help function, ● Feedback form. ● Log-in ● Email address for web master, organization contact
Increase fundraising (ISTS, 2003; Weimann, 2004)	<ul style="list-style-type: none"> ● Publicizing need for funds ● Providing options for collecting funds 	<ul style="list-style-type: none"> ● Payment instruction & facility ● E-commerce application ● Hyperlinks to other resources
Diffuse propaganda (ISTS, 2004; Weimann, 2004)	<ul style="list-style-type: none"> ● Posting resources in multiple languages. ● Providing links to forums, videos & other groups' web sites ● Using web sites as an online clearinghouses for statements from leaders 	<ul style="list-style-type: none"> ● Content management ● Hyperlinks ● Directory for documents ● Navigation support ● Search, browsable index ● Free web site hosting ● Accessible
Increase publicity (Coll and Glasser, 2005; Jenkins, 2004)	<ul style="list-style-type: none"> ● Advertising groups' events, martyrs, history, ideologies. ● Providing groups' interpretation of the news 	<ul style="list-style-type: none"> ● Downloadable files ● Animated & flashy banner, logo, slogan ● Clickable maps, ● Information resources (e.g., international news)
Overcome obstacles from law enforcement & military (Coll and Glasser, 2005; Kelley, 2001)	<ul style="list-style-type: none"> ● Send encrypted messages via email, forums, or post on web sites ● Move web sites to different servers so that they are protected 	<ul style="list-style-type: none"> ● Anonymous email accounts ● Password protected or encrypted services ● Downloadable encryption software ● Email security ● Stenography
Provide recruitment & training (ISTS, 2003; Weimann, 2004)	<ul style="list-style-type: none"> ● Hosting martyrs stories, speeches, multimedia that are used for recruitment. ● Using flashy logos, banners, cartoons to appeal to sympathizers with specialized skills & similar views ● Build massive & dynamic online libraries of training resources 	<ul style="list-style-type: none"> ● Interactive services (e.g., games, cartoons, maps) ● Online registration process ● Directory ● Multimedia (e.g., videos, audios, images) ● FAQ, alerts ● Virtual community

limited the comprehensiveness of their analyses. Furthermore, none of the studies have provided empirical evidence of the levels of technical sophistication or compared terrorist/extremist organizations' cyber capabilities with those of mainstream organizations. Since technical knowledge required to maintain Web sites provides an indication of terrorist/extremist organizations' technology adoption strategies (Jackson, 2001), we believe it is important to analyse the technologies required to maintain terrorist/extremists' Web sites from the perspectives of technical sophistication, content richness, and Web interactivity.

2.3. Dark Web collection building

The first step towards studying the terrorist/extremist Web presence is to capture terrorist Web sites and store

them in a repository for further analysis. Web collection building is the process of gathering and organizing unstructured information from pages and data on the Web. Previous studies have suggested three types of approaches to collecting Web contents in specific domains: manual approach, automatic approach, and semiautomatic approach.

In order to build the September 11 and Election 2002 Web Archives (Schneider et al., 2003), the Library of Congress collected seed URLs for a given theme. The seeds and their close neighbors (distance 1) are then downloaded. The limitation of such a manual approach is that it is time-consuming and inefficient.

Anderson (2003) used an automatic approach in the "Paradigma" project. The goal of Paradigma is to archive Norwegian legal deposit documents on the Web. It

Table 2
Organizations that capture and analyze terrorists' web sites

Organization	Description	Access
<i>Archive</i>		
1. Internet Archive (IA)	1996—Collect open access HTML pages (every 2 months)	Via http://www.archive.org
<i>Research Center</i>		
2. Anti-terrorism Coalition (ATC)	2003—Jihad Watch. Has 448 terrorist Web sites & forums	Via http://www.atcoalition.net
3. Artificial Intelligence (AI) Lab, University of Arizona	2003—Spidering (every 2 months) to collect terrorist Web sites. Has 1000s Web sites: US Domestic, Latin America, & Middle Eastern Web sites	Via testbed portal called Dark Web Portal
4. MEMRI	2003—Jihad & Terrorism Studies Project.	Access reports via http://www.memri.org
5. Site Institute	2003—Capture Web sites every 24 hrs. Extensive collection of 1000s of files.	Access reports & fee-based intelligence services http://siteinstitute.org
6. Weimann (Univ. Haifa, Israel)	1998—Capture Web sites daily. Extensive collection of 1000s of files.	Closed collection
<i>Vigilante community</i>		
7. Internet Haganah	2001—Confronting the Global Jihad Project. Has 100s links to Web sites.	Provides snapshots of terrorist Web sites http://haganah.us

employed a focused Web crawler (Chakrabarti et al., 1999), an automatic program that discovers and downloads Web sites in particular domains by following Web links found in the HTML pages of a starting set of WebPages. Metadata was then extracted and used to rank the Web sites in terms of relevance. The automatic approach is more efficient than the manual approach; however, due to the limitations of current focused crawling techniques, automatic approaches often introduce noise (off-topic Web pages) into the collection.

The “Political Communications Web Archiving” group employed a semiautomatic approach to collecting domain-specific Web sites (Reilly et al., 2003). Domain experts provided seed URLs as well as typologies for constructing metadata that can be used in the crawling process. Their project’s goal is to develop a methodology for constructing an archive of broad-spectrum political communications over the Web. We believe that the semiautomatic approach is most suitable for collecting terrorist/extremist Web sites because it combines the high accuracy and high efficiency of manual and automatic approaches.

2.4. Dark Web content analysis

In order to reach an understanding of the various facets of terrorist/extremist Web usage and communications, a systematic analysis of the Web sites’ content is required. Researchers in the terrorism domain have used observation and content analysis to analyse Web site data. In Bunt’s (2003) overview of Jihadi movements’ presence on the Web, he described the reaction of the global Muslim community to the content of Jihadi terrorist Web sites. His assessment of the influence such content had on Muslims and Westerners was based on a qualitative analysis of

message contents extracted from Taliban and Al-Qaeda Web sites. Tsftati and Weimann (2002) conducted a content analysis of the characteristics of terrorist groups’ communications. They said that the small size of their collection and the descriptive nature of their research questions made a quantitative analysis infeasible.

Demchak et al. (2001) provided a well-defined methodology for analyzing communicative content in government Web sites. Their work focused on measuring “openness” of government Web sites. To achieve this goal they developed a Web site Attribute System (WAES) tool that is basically composed of a set of high level attributes such as transparency and interactivity. Each high level attribute is associated with a second layer of attributes at a more refined level of granularity. For example, the increase of “operational information” and “responses” on a given Webpage can induce an increase in the openness level of a government Web sites. This WAES system is an example of a well-structured and systematic content analysis methodology.

Demchak and Friis’ work provides guidance for the present study. However, the “openness” attributes used in their work were designed specifically for e-Government studies. We surveyed research in e-Commerce, e-Government, and e-Education domains and identified several sets of attributes that could be used to study the technical advancement and effectiveness of terrorists/extremists’ use of the Internet.

Palmer and Griffith’s (1998) study identified a set of 15 attributes (called “technical characteristics” in the original work) to evaluate two aspects of e-Commerce Web sites: technical sophistication and media richness. More specifically, the technical sophistication attributes measures the level of advancement of the techniques used in the design of

Web sites. For example, “use of HTML frames,” “use of Java scripts,” etc. The media richness attributes measure how well the Web sites use multimedia to deliver information to their users, e.g., “hyperlinks,” “images,” “video/audio files,” etc.

Another set of attributes called Web interactivity has been widely adopted by researchers in e-Government and e-Education domains to evaluate how well Web sites facilitate the communications among Web site owners and users. Two organizations, the United Nations Online Network in Public Administration and Finance (UNPAN; www.unpan.org) and the European Commission’s IST program (www.cordis.lu/ist/) have conducted large-scale studies to evaluate the interactivity of government Web sites of major countries in the world. The Web interactivity attributes can be summarized into three categories: one-to-one-level interactivity, community-level interactivity, and transaction-level interactivity.

The one-to-one-level interactivity attributes measure how well the Web sites support individual users to give feedback to the Web site owners (e.g., provide email contact, provide guest book functions, etc.). The community-level interactivity attributes measure how well the Web sites support the two-way interaction between site owners and multiple users (e.g., use of forums, online chat rooms, etc.). The transaction-level interactivity measures how well users are allowed to finish tasks electronically on the Web sites (e.g., online purchasing, online donation, etc.).

Chou’s (2003) study proposed a detailed four-level framework to analyse e-Education Web site’s level of advancement and effectiveness. Attributes in the first level (called learner-interface interaction) of Chou’s framework are very similar to the technical sophistication attributes used in Palmer and Griffith’s (1998) study. Attributes in the other three levels (learner–content interaction, learner–instructor interaction, and learner–learner interaction) of Chou’s framework are similar to the three-level Web interactivity attributes used in the e-Government evaluation projects as mentioned above.

To date, no study has employed the technical sophistication, media richness, and Web interactivity attributes as well as the WAES framework in the terrorism domain. We believe that these Web content analysis metrics can be applied in terrorist/extremist Web site analysis to deepen our understanding of the terrorist’s tactical use of the Web.

3. Proposed methodology: Dark Web collection and analysis

The research questions postulated in our study are:

- (1) What design features and attributes are necessary to build a highly relevant and comprehensive Dark Web collection for intelligence and analysis purposes?
- (2) For terrorist/extremist Web sites, what are the levels of technical sophistication in their system design?
- (3) For terrorist/extremist Web sites, what are the levels of richness in their online content?

- (4) For terrorist/extremist Web sites, what are the levels of Web interactivity to support individual, community, and transaction interactions?

To study the research questions, we propose a Dark Web analysis tool which contains several components: a systematic procedure for collecting and monitoring Dark Web contents and a Dark Web Attribute System (DWAS) to enable quantitative analysis of Dark Web content (see Fig. 1).

3.1. Dark Web collection building

The first step towards studying terrorists’ tactical use of the Web is to build a high-quality Dark Web collection. To ensure the quality of our collection, based on our review of Web collection building methodologies, we propose to use a semi-automated approach to collecting Dark Web contents (Reid et al., 2004). Our collection build approach contains the following steps (see Fig. 2).

(1) *Identify terrorist/extremist groups*: Defining terrorism is complicated by the fact that people almost never define themselves as terrorists and the use of the label by others often has political overtones. We start the collection building process by identifying the groups that are considered by authoritative sources as terrorist/extremist groups. The sources include government agency reports (e.g., US State Department reports, FBI reports, government reports from United Kingdom, Australia, Japan, and P. R. China, etc.), authoritative organization reports (e.g., Counter-Terrorism Committee of the UN Security Council, US Committee for A Free Lebanon, etc.), and studies published by terrorism research centers such as the Anti-Terrorism Coalition (ATC), the Middle East Media Research Institute (MEMRI), Dartmouth College, etc. Information such as terrorist group names, leaders’ names, and terrorist jargons are identified from the sources to create a terrorism keyword lexicon for use in the next step.

(2) *Identify terrorist/extremist group URLs*: We manually identify a set of seed terrorist group URLs from two sources. First, terrorist group URLs can be directly identified from the authoritative sources and literatures used in the first step. Second, terrorist group URLs can be identified by using the terrorism keyword lexicon to query major search engines on the Web. The identified set of terrorist group URLs will serve as the seed URLs for the next step.

(3) *Expand terrorist/extremist URL set through link and forum analysis*: After identifying the seed URLs, out-links and in-links of the seed URLs were automatically extracted using link-analysis programs. The out-links are extracted from the HTML contents of “favorite link” pages under the seed Web sites. The in-links are extracted from Google in-link search service through Google API. Automatic out-link and in-link expansion is an effective way to expand the scope of our collection. We also have language experts who browse the contents of terrorist supporting forums and

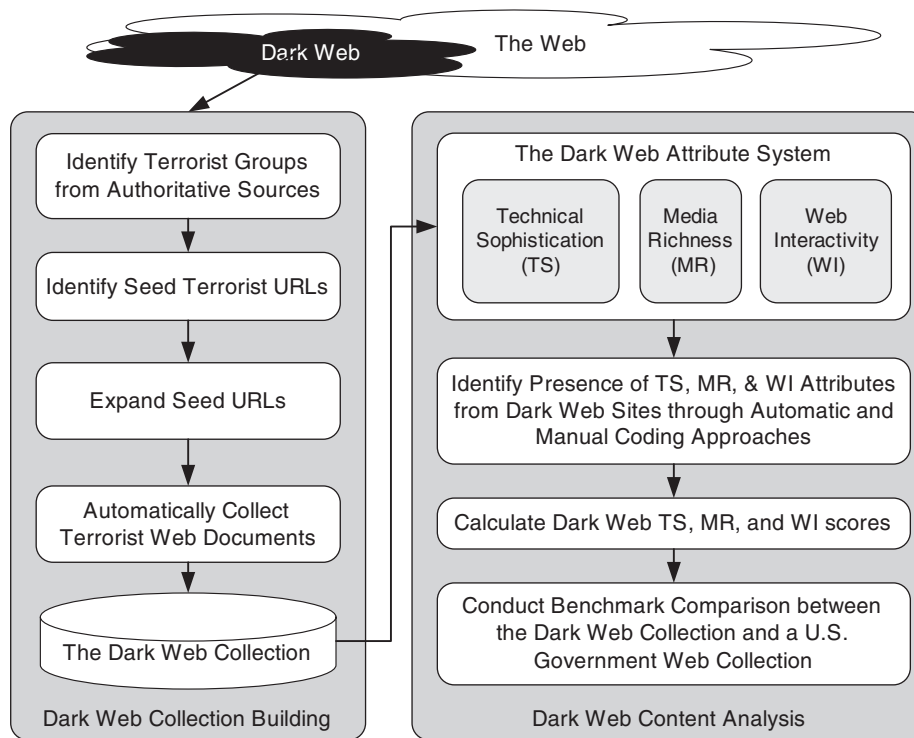


Fig. 1. The Dark Web collection building and content analysis framework.

extract the terrorist/extremist URLs posted by terrorist supporters. Because bogus or unrelated Web sites can make their way into our collection through the expansion, we have developed a robust filtering process based on evidence and clues from the Web sites. Aside from sites which explicitly identify themselves as the official sites of a terrorist organization or one of its members, a Web site that contains even minor praise of or adopts ideologies espoused by a terrorist group is included in our collection.

(4) *Download terrorist/extremist Web site contents*: Once the terrorist/extremist Web sites are identified, a program is used to automatically download all their contents. Unlike the tools used in previous studies, our program was designed to download not only the textual files (e.g., HTML, TXT, PDF, etc.) but also multimedia files (e.g., images, video, audio, etc.) and dynamically generated Web files (e.g., PHP, ASP, JSP, etc.). Moreover, because terrorist organizations set up forums within their Web sites whose contents are of special value to research communities, our program also can automatically log into the forums and download the dynamic forum contents. The automatic downloading method allows us to effectively build Dark Web collections with millions of documents. This would greatly increase the comprehensiveness of our Dark Web study.

To keep the Dark Web collection comprehensive and up-to-date, Steps 2 to 4 are periodically repeated. Collections built using such a recursive procedure can also provide information about the evolution and diffusion of the Dark Web.

3.2. Dark Web content analysis: the Dark Web Attribute System (DWAS)

Instead of using observation-based qualitative analysis approaches (Thomas, 2003); we propose a systematic approach to enable the quantitative study of terrorist/extremist groups' use of the Web. The proposed DWAS is similar to the WAES framework in Demchak et al.'s study (2001). However, instead of the openness attributes used in WAES, our framework focuses on the attributes that could help us better understand the level of advancement and effectiveness of terrorists' Web usage, namely, technical sophistication attributes, content richness attributes (an extension of the traditional media richness attributes), and Web interactivity attributes. Based on previous literatures in e-Commerce (Palmer and Griffith, 1998), e-Government (Demchak et al., 2001), and e-Education domains (Chou, 2003; Hillman et al., 1994), we selected 13 technical sophistication attributes, five content richness attributes, and 11 Web interactivity attributes for our DWAS framework. A list of these attributes is summarized in Tables 3a–c.

(1) *Technical sophistication (TS) attributes*: The technical sophistication attributes can be grouped into four categories as shown in Table 3a. The first category of four attributes, called the basic HTML technique attributes, measures how well the basic HTML layout techniques (i.e., Lists, Tables, Frames, and Forms) are applied in Web sites to organize Web contents. The second category, called the embedded media attributes, measures

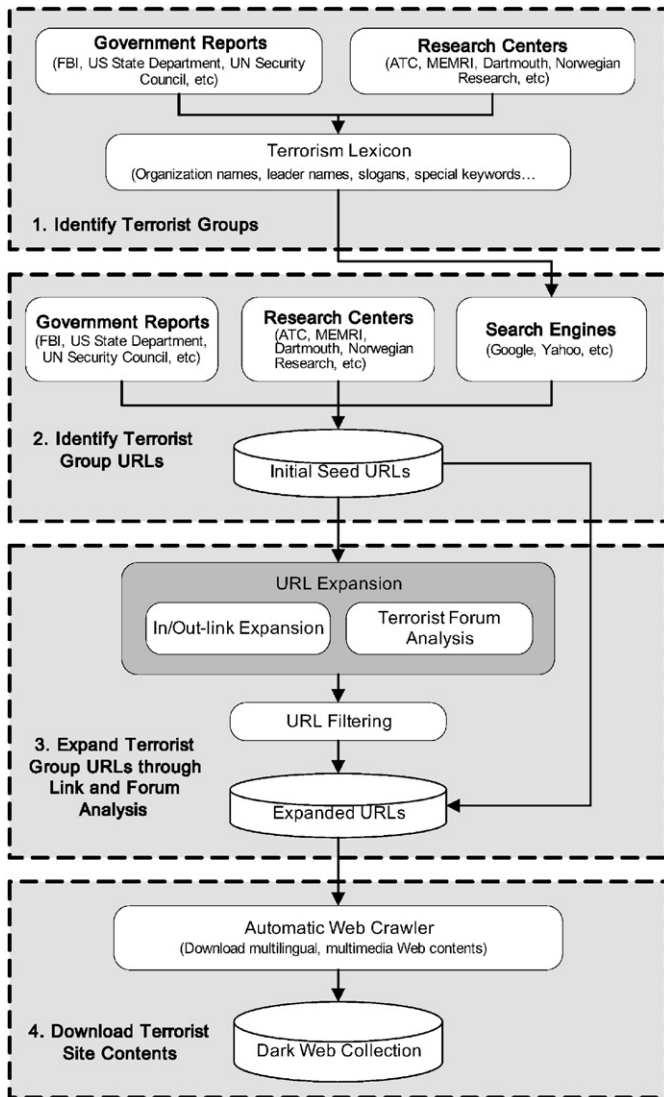


Fig. 2. The Dark Web Collection Building Approach.

how well the Web sites deliver their information to the user in multimedia formats such as images, animations, and audio/video clips. The third category of three attributes, called the advanced HTML attributes, measures how well advanced HTML techniques such as DHTML, SHTML, predefined and self-defined script functions (e.g., JavaScript, VBScript, etc.) are applied to implement security and dynamic functionalities. The last category, called the dynamic Web programming attributes, measures how well dynamic Web programming languages such as PHP, ASP, and JSP are utilized to implement dynamic interaction functionalities such as user login, online request or application, and online transaction processing. The four technical sophistication attributes and associated sub-attributes are present in most of the Dark Web sites we collected.

The presence of different attributes indicates different level of technical sophistication. For example, a Web site

Table 3a
Technical sophistication attributes

TS attributes	Weights
Basic HTML techniques	
Use of lists	1
Use of tables	2
Use of frames	2
Use of forms	1.5
Embedded multimedia	
Use of background image	1
Use of background music	2
Use of stream audio/video	3.5
Advanced HTML	
Use of DHTML/SHTML	2.5
Use of predefined script functions	2
Use of self-defined script functions	4.5
Dynamic web programming	
Use of CGI	2.5
Use of PHP	4.5
Use of JSP/ASP	5.5

Table 3b
Content richness attributes

CR attributes	Scores
Hyperlink	No. of hyperlinks
File/Software download	No. of downloadable documents
Image	No. of images
Video/audio file	No. of video/audio files

Table 3c
Web interactivity attributes

WI attributes	Weights
<i>One-to-one interactivity</i>	
Email feedback	1.75
Email list	2.25
Contact address	1.25
Feedback Form	2.75
Guest book	1.5
<i>Community-level interactivity</i>	
Private message	4.25
Online forum	4.25
chat room	4.75
<i>Transaction-level Interactivity</i>	
Online shop	4
Online payment	4
Online application form	4

which uses JSP techniques should be considered more technically sophisticated than a site which only uses static HTML. Different weights should be assigned to the attributes to reflect the differences (Chou, 2003). We determined the weights based on Web experts' opinions

collected through an email survey. Surveys were sent to Web masters and network administrators of several Web sites belonging to the University of Arizona and they were encouraged to forward the survey to their Web master colleagues. In the survey, we asked the experts to give each of our attributes a weight of 1–10 (1 is the least advanced/sophisticated). Six experts sent their responses back to us. For each attribute, the average weight assigned by the experts was used in the final framework. Among the six experts, two are Web masters of academia Web sites, two are Web masters of commercial Web sites, one is a Web developer in a commercial company, and the last one is a professor teaching Web development courses in a university. On average, they have seven years of professional experience in Web technology. To ensure the reliability of the weights, we conducted reliability test on the experts' answers. The reliability score (Cronbach's alpha) calculated for the experts' answers was 0.89 which was well above the 0.70 required for acceptable scale reliability (Nunnally, 1978). The TS attributes and their weights are summarized in Tables 3a.

(2) *Content richness (CR) attributes*: In traditional media richness studies, researchers only focused on the variety of media used to deliver information (Trevino et al., 1987; Palmer and Griffith, 1998). However, to have a deep understanding of the richness of Dark Web contents, we would like to measure not only the variety of the media but also the amount of information delivered by each type of media. In our study, we expand the media richness concept by taking the volume of information into consideration. More specifically, as shown in Table 3b, we calculated the average number of four types of Web elements: hyperlinks, downloadable documents, images, and video/audio files, as the indication of Dark Web content richness.

(3) *Web interactivity (WI) attributes*: For the Web interactivity attributes (see Table 3c), we followed the standard built by the UNPAN and the European Commission's IST program as well as Chou's (2003) work to group the attributes into three levels: the one-to-one-level interactivity, the community-level interactivity, and the transaction level interactivity. The one-to-one-level interactivity contains five attributes (i.e., Email Feedback, Email List, Contact Address, Feedback Form, and Guest Book) that provide basic one-to-one communication channels for Dark Web users to contact the terrorist Web site owners (see Table 3c). The community-level interactivity contains three attributes (i.e., Private Message, Online Forum, and Chat Room) that allow Dark Web site owners and users to engage in synchronized many-to-many communications with each other. The transaction-level interactivity contains three attributes (i.e., Online Shop, Online Payment, and Online Application Form) that allow Dark Web users to complete tasks such as donating to terrorist/extremist groups, applying for group membership, etc. The presence of these attributes in the Dark Web sites indicates how well terrorists/extremists utilize Internet

technology to facilitate their communications with their supporters.

Similar to the TS attributes, different weights should be assigned to the WI attributes to indicate their different levels of support on communications. We asked Web experts to assign weights of 1 to 10 to the WI attributes in the same email survey where the TS attributes weights were determined. The WI attributes and their weights are summarized in Table 3c.

We developed strategies to efficiently and accurately identify the presence of the DWAS attributes from Dark Web sites. The TS and CR attributes are marked by HTML tags in page contents or file extension names in the page URL strings. For example, an HTML tag "<image>" indicates that an image is inserted into the page content. A URL string ending with ".jsp" indicates that the page utilizes JSP technology. We developed programs to automatically analyse Dark Web page contents and URL strings to extract the presence of the TS and CR attributes. Since there are no clear indications or rules that a program could follow to identify WI attributes from Dark Web contents with a high degree of accuracy, we developed a set of coding scheme to allow human coders to identify their presence in Dark Web sites. Technical sophistication, content richness, and Web interactivity scores are calculated for each Web site based on the presence of the attributes to indicate how advanced and effective the site is in terms of supporting terrorist/extremist groups' communications and interactions.

4. Case study

To test our proposed approach, we conducted a case study to collect and analyse the Web presence of major Middle Eastern terrorist groups. We also conducted a benchmark comparison between the terrorist/extremist Web sites and US federal and state government Web sites to evaluate the terrorist/extremist organizations' online capabilities. The terrorist/extremist groups we studied mainly include Islamic terrorist groups rooted in Middle Eastern countries, for example, Al Qaeda, Palestinian Islamic Jihad, Hamas, etc. These terrorist/extremist groups are the focus of most current counter-terrorism studies. We chose US government Web sites as benchmarks because government Web sites and terrorist/extremist Web sites have common overall objectives—to inform the public about their goals, programs, and strategies. To achieve this objective, similar Web features must be implemented in both government and terrorist/extremist Web sites. Furthermore, the US government was ranked the top in the world by the CyPRG group (<http://www.cyprg.arizona.edu/>) in terms of Web technical sophistication and interactivity. With the US government Web sites as high-standard benchmarks, we can better understand the terrorist/extremist Web sites' levels of technical advancement and effectiveness.

4.1. Building Dark Web research testbed

Following the collection building procedure discussed in Section 3.1, we created a Middle Eastern terrorist/extremist Web site collection and a US government Web site collection as the testbeds for this study.

The Middle Eastern terrorist/extremist Web collection was created in June of 2004. We identified 36 Middle Eastern terrorist/extremist groups from authoritative sources mentioned in Section 3.1. Based on the information of these terrorist/extremist groups, we constructed a lexicon of Middle Eastern terrorism keywords with the help of Arabic language experts. Examples of relevant keywords include terrorist leaders' names such as "لادن بن المجاهد الشيخ" (Sheikh Mujahid bin Laden); Terrorist groups' names such as "خلق ايران" ("Khalq Iran"), and special words used by terrorists/extremists such as "حرب صليبية" ("Crusader's War") and "الكفار" ("Infidels"). This lexicon was used to query major search engines for identification and retrieval of terrorist/extremist groups' URLs. The URLs identified from the search engines, together with the terrorist/extremist URLs listed in the terrorism literature and reports, served as seed URLs for the out-link and in-link expansion process. We performed a one level deep in-link expansion using Google's in-link search tool and a level deep out-link expansion. After carefully filtering the expansion results, we obtained the URLs of 86 Middle-Eastern terrorist/extremist Web sites. Using SpidersRUs, a digital library building toolkit developed by our group, we collected about 222,000 multimedia Web documents from the identified terrorist/extremist Web sites.

Table 4 summarizes the detailed file type breakdown of the terrorist/extremist collection. 179,223 out of the total 222,687 documents in the terrorist/extremist collection are indexable files. These are textual files such HTML files, plain text files, PDF/Word documents, and dynamic files

generated by Web applications (e.g., ASP, JSP, etc.). Interestingly, the majority of indexable files (130,972 files out of 179,223 total files) in the terrorist/extremist collection are dynamic files. We conducted a preliminary analysis on the contents of these dynamic files and found that most dynamic files were forum postings. This indicates that online forums play an important role in terrorists/extremists' Web usage. Other than indexable files, multimedia files also make a significant presence in the terrorist/extremist collection. While the quantity of multimedia files is not as large as the indexable files, multimedia files are the largest category in the collection in terms of their volume. This indicates heavy use of multimedia technologies in terrorist/extremist Web sites. The last two categories, archive files (1281 files) and non-standard files (7019 files), made up less than 5% of the collection. Archive files are compressed file packages such .zip files and .rar files. They could be password-protected. Non-standard files are files that cannot be recognized by the Windows operating system. These files may be of special interest of terrorism researchers and experts because they could be encrypted information created by terrorists/extremists. Further analysis is needed to study the contents of these two types of files.

The benchmark US government Web collection was built in July of 2004. All 92 federal and state government URLs under Yahoo! "Government" category were selected as seed URLs. Around 277,000 Web documents were automatically collected from these government Web sites using the SpidersRUs toolkit. The detailed file type breakdown of the US government Web collection is summarized in Table 5. The file type distribution of the government collection is similar to the terrorist/extremist collection. Indexable files (221,684 files) are the largest category, majority of which are dynamic files (145,590 files). However, in the government collection, we did not find as many forum postings as in the terrorist/extremist

Table 4
Middle-eastern terrorist/extremist web collection file types

Terrorist/extremist collection	No. of files	Volume (bytes)
Grand total	222,687	12,362,050,865
Indexable files total	179,223	4,854,971,043
HTML files	44,334	1,137,725,685
Word files	278	16,371,586
PDF files	3145	542,061,545
Dynamic files	130,972	3,106,537,495
Text files	390	45,982,886
Powerpoint files	6	6,087,168
XML files	98	204,678
Multimedia files total	35,164	5,915,442,276
Image files	31,691	525,986,847
Audio files	2554	3,750,390,404
Video files	919	1,230,046,468
Archive files	1281	483,138,149
Non-standard files	7019	1,108,499,397

Table 5
US government web collection file types

US Government collection	No. of files	Volume (bytes)
Grand total	277,274	19,341,345,384
Indexable files total	221,684	6,502,288,302
HTML files	71,518	2,632,912,620
Word files	298	210,906,045
PDF files	841	663,293,376
Dynamic files	145,590	2,071,734,849
Text files	2878	555,403,447
Excel files	4	98,560
Powerpoint files	5	725,017
XML files	554	367,214,389
Multimedia files total	49,582	10,835,029,216
Image files	45,707	850,011,712
Audio files	3429	8,153,419,931
Video files	449	1,831,597,573
Archive files	538	286,312,990
Non-standard files	5471	1,717,714,876

collection. Many dynamic files in the government collection are articles dynamically retrieved from large document database on users' requests. Multimedia files also have a significant presence in the government collection, indicating heavy multimedia usage in government Web sites.

4.2. Collection analysis and benchmark comparison

Following the DWAS approach, presence of the technical sophistication and media richness attributes was automatically extracted from the collections using programs. Presence of the Web interactivity attributes was extracted from each Web site by language experts based on the coding scheme in DWAS. Because of the time limitation, language experts only examined the top two level Web pages in each Web site. For each Web site in the two collections, three scores (technical sophistication, content richness, and Web interactivity) were calculated based on the presence of the attributes and their corresponding weights in DWAS. Statistical analysis was conducted to compare the advancement/effectiveness scores achieved by the terrorist/extremist collection and the US government collection.

4.2.1. Benchmark comparison results: technical sophistication

The technical sophistication comparison results are shown in Table 6. The results showed that:

- The US government Web sites are significantly more advanced than the terrorist Web sites in terms of basic HTML techniques ($p < 0.0001$). Government agencies paid much attention to the design of their Web sites and they used many of the HTML features to organize their Web contents. Terrorists/extremists, on the other hand, did not organize the contents on their Web sites very well.
- The US government Web sites are significantly more advanced than the terrorist Web sites in terms of utilizing dynamic Web programming languages ($p = 0.0066$). Most government Web sites employed Web programming technologies (e.g. PHP, ASP, JSP, etc.) to implement functionalities such as user login, online application, online purchase, etc. Few terrorist/extremist Web sites implemented such dynamic functionalities.

Table 6
Technical sophistication comparison results

TS attributes	Weighted average score		<i>t</i> -Test result
	US	Terrorists	
Basic HTML Techniques	0.9130434	0.710526	$p < 0.0001^{**}$
Embedded Multimedia	0.565217	0.833333	$p = 0.0027^{**}$
Advanced HTML	1.789855	1.771929	$p = 0.139$
Dynamic Web Programming	2.159420	1.407894	$p = 0.0066^{**}$
Average	1.356884	1.180921	$p = 0.06$

** Significant level is at 0.05.

- There is no significant difference between the terrorist Web sites and the US government Web sites in terms of applying advanced HTML techniques at a significant level of 0.05 ($p = 0.139$).
- The terrorist Web sites have a significantly higher level of embedded media usage than the US government Web sites ($p = 0.0027$). This unique characteristic of terrorist/extremist Web sites is discussed in detail below.
- When taking all four sets of attributes into consideration, there is no significant difference between the technical sophistication of the Middle-Eastern terrorist Web sites and the US government Web sites at a significant level of 0.05 ($p = 0.06$).

The extensive use of media in terrorist/extremist groups' Web sites is of special interest. While the terrorist/extremist groups' are not as good as the US government in terms of organizing their Web pages into clear layouts or implementing dynamic Web functionalities, they employed a significantly higher level of embedded multimedia techniques, especially images and audio/video clips, to catch the interests of their target audience. In the terrorist/extremist groups' collection, 46% of the Web sites embedded audio/video clips into their pages, while only 29% of the US government Web sites provided audio/video clips.

Multimedia content is more attractive and tends to leave a stronger impression on people than pure textual content. For example, militant Islamic group Hamas foments a violent resistance to their "enemies" by disseminating graphic posters on their Web sites (see Fig. 3). Moreover, terrorists often post images, audio, or video clips from their leaders or martyrs to boost the spirit of their members and supporters. For example, Osama bin Laden's portrait appears in homepages of many Middle Eastern terrorist/extremist Web sites. Recently, posters of the Iraqi terrorist leader Abu Mus'ab Zarqawi who is suspected to be responsible for the beheading of several western hostages can also be found in Middle-Eastern terrorist Web sites (see Fig. 4). These posters explicitly mention that Abu Mus'ab Zarqawi is a "beheader" and praise his brutal



Fig. 3. A Hamas poster inviting men to join their military struggle. The text on the poster says "Have you fought for the sake of God? You say no. Then you should have your mouth shot." Source: <http://www.palestine-info.com>.



Fig. 4. A poster depicting terrorist leader in Iraq, Abu Mus’ab Zaraqawi. The text on the poster says “Emir Zaraqawi, may God save him. Eagle of Iraq, volcano of Jihad, and the beheader.” Source: <http://www.islamic-f.net/vb/>.

also found that an Iraqi terrorist/extremist group posted pictures of executed “traitors” on their Web sites, warning other Iraqi people not to cooperate with the US Forces. Materials of such nature are usually considered to be too shocking to televise by most TV news producers. However through the Internet, terrorists/extremists have successfully spread these gruesome materials to as many people as possible, especially in the West where Internet use is more common.

4.2.2. Benchmark comparison results: content richness

The content richness comparison results are summarized in Table 7. The results showed that:

- The US government Web sites provided significantly more hyperlinks ($p < 0.0001$), downloadable documents ($p = 0.0103$), and video/audio clips ($p < 0.0001$) than the terrorist/extremist Web sites.
- The US government Web sites provided more images than the terrorist/extremist Web sites; but the difference is not significant at a significant level of 0.05.
- Overall, the terrorist/extremist Web sites are not as good as the US government Web sites in terms of content richness ($p < 0.0001$) because the volumes of contents in terrorist/extremist Web sites are often smaller than US government Web sites.

The content richness comparison results are not contradictory with the technical sophistication comparison results. The content richness results showed that the US government Web sites provide a larger volume of multimedia content; while the technical sophistication results indicated that a higher percentage of terrorist/extremist groups’ Web sites provide multimedia contents. The terrorist/extremist Web sites also utilize more advanced technology to deliver their multimedia contents.

One possible explanation for the smaller volume of multimedia content provided by the terrorist/extremist groups’ Web sites is the lower capacity and instability of terrorists/extremists’ Web servers. Unlike the US government Web sites which are usually hosted on dedicated Web servers, many of the terrorist/extremist groups’ Web sites in our collection are hosted on Web servers provided by free public ISPs such as Geocities. The public Web servers usually have restrictions on the size and bandwidth of the



Fig. 5. A list of audio clips from the Web site of extremist cleric sheikh Hamed Al Ali which consists of preaching in the Salafi ideology and political issues. Source: <http://www.h-alali.net>.

killing of innocents as a way to protect Iraq. Terrorists/extremists also post images and audio/video clips of their “martyrdom operations” as a way to demonstrate their resolve to fight their enemies and inspire their supporters. Many movie clips of several suicide bombing attacks in Iraq were posted by terrorists in one of the terrorist online forums (<http://www.lb.dm.net.lb/ubb/Forum4/>) to show off their “triumph over the US invaders.” The “Fighting Islamic Group” guerilla posted a set of detailed documentations with pictures describing their assassination attempt of Libyan president Mu’amar Kadhafi and praising the “heroism” of their members (see Fig. 5).

The multimedia content posted on terrorist/extremist Web sites is not only for terrorist supporters but for enemies. For example, the video clip of American Nicholas Berg being beheaded was spread to the public from a Malaysian terrorist Web site. The video of the final minutes of another American hostage, Robert Jacobs, was first posted on Middle Eastern militant group’s Web sites. We

Table 7
Content richness comparison results

CR attributes	Average counts per sites		t-Test result
	US	Terrorists	
Hyperlink	3513.254654	3172.658483	$p < 0.0001^{**}$
Downloaded documents	400.9674532	151.868427	$p = 0.0103^{**}$
Image	582.352456	540.0484563	$p = 0.466$
Video/audio file	91.55434783	50.9736828	$p < 0.0001^{**}$

** Significant level is at 0.05.

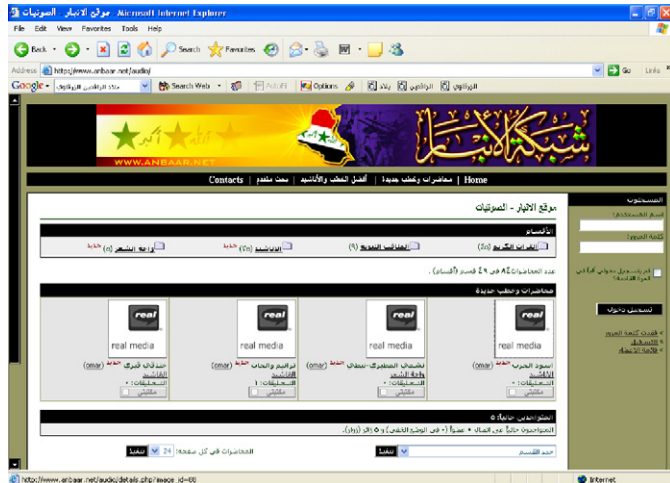


Fig. 6. “Holy war” songs and hymns presented on Anbaar Iraqi terrorist/extremist group’s website audio section. Source: <http://www.anbaar.net/audio>.

Web sites they host. The restrictions would limit terrorist/extremist groups’ ability to host multimedia information on their Web sites. The instability of the terrorist/extremist groups’ Web sites also makes it difficult for them to host multimedia information. Many Web sites frequently move their Web contents to other Web servers because their old sites were shut down by ISPs or hacked. While textual Web pages can be quickly and easily duplicated to the new servers, multimedia documents are more difficult to transfer and more prone to loss because of their larger size.

Nevertheless, terrorist/extremist groups still manage to host a considerable amount of downloadable documents and multimedia information on their Web sites. These media cover a wide variety of topics ranging from propaganda campaigns to tutorials of weapon operations and guerilla tactics. For example, the Web site of extremist cleric sheikh Hamed Al Ali (see Fig. 5) hosts a list of audio clips consisting of preaching in the Salafi ideology and political issues. The Anbaar Iraqi terrorist/extremist group’s Web site (see Fig. 6) provides a collection of songs and hymns praising the “Holy war” that they are conducting.

4.2.3. Benchmark comparison results: Web interactivity

Table 8 summarizes the Web interactivity comparison results. The results showed that:

- In terms of supporting one-to-one-level interactivity, the US government agencies are doing significantly better than terrorist/extremist Web sites by providing their contact information (e.g., email, mail address, etc.) on their sites ($p = 0.024$). Because of their covert nature, terrorist/extremist groups seldom disclose their contact information on their Web sites.
- In terms of support community-level interactivity, terrorist/extremist Web sites are doing significantly better than government Web sites by providing online

Table 8
Web interactivity comparison results

WI attributes	Weighted average score		t-Test result
	US	Terrorist	
One-to-one	0.342857	0.292169	$p = 0.024^{**}$
Community	0.028571	0.168675	$p = 0.0025^{**}$
Transaction	0.3	Not presented	
Average (transaction not included)	0.185714	0.230422	$p = 0.056$

** Significant level is at 0.05.

forums and chat rooms ($p = 0.0025$). Few government agencies provided such online forum and chat room support on their Web sites.

- Our experts did not identify transaction-level interactivity in terrorist/extremist Web sites, although such interactivity might be hidden in their sites.
- Taking both one-to-one and community level interactivity into consideration, we did not find significant difference between the terrorist/extremist Web sites and the US government Web sites ($p = 0.056$) at a significant level of 0.05.

Several previous studies implied that terrorists are relying on Internet-based communication tools such as online chat rooms and forums to facilitate their daily communication, command and control, and even operation planning and coordination (Zhou et al., 2005; Whine, 1999; FBIS, 1995). Our results further confirmed these observations. The Middle Eastern terrorist/extremist groups are very active in terms of hosting and maintaining online forums and bulletin boards. Among the largest terrorist-supporting forum that we have been monitoring, www.shawati.com has 31,894 registered forum members and 418,196 posts; www.kuwaitchat.net has 11,531 registered members and 624,694 posts. Not all of the forum members are terrorists or extremists. Many of them are just supporters or sympathizers. Members of these large forums participate in daily discussions, express their support of the terrorist groups, and reinforce each other’s beliefs in the terrorist/extremist groups’ courses. They sometimes can get messages directly from active members of terrorist/extremist groups. For example, messages from the Iraqi terrorist leader, Abu Mus’ab Zarqawi can often be found in online forum www.islamic-f.net (see Fig. 7). These dynamic forums provide snapshots of terrorist/extremist groups’ activities, communications, ideologies, relationships, and evolutionary developments.

5. Conclusions and future directions

In this study, we proposed a systematic procedure to collect Dark Web contents and a Dark Web Attribute System (DWAS) to enable quantitative analysis of terrorists’ tactical use of the Internet. The automatic

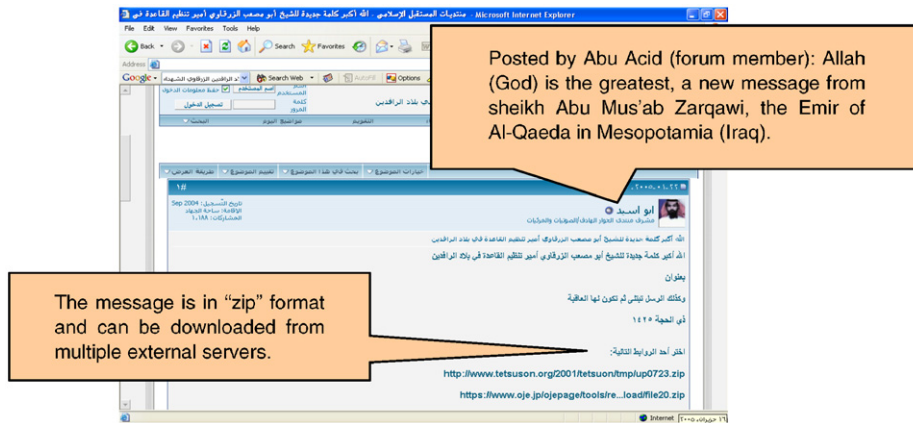


Fig. 7. Discussion forums are used to share important messages form terrorist leaders among the members of the terrorist groups and their supporters. Source: www.islamic-f.net.

collection building and content analysis components used in the proposed methodology allow the efficient collection and analysis of thousands of Dark Web documents. This enables our Dark Web study to achieve a high level of comprehensiveness than previous manual approaches. Furthermore, the DWAS is a systematic content analysis tool that, we believe, brings more insights into terrorist/extremist groups' Internet usages than previous observation-based studies provided.

Using the proposed collection building procedure and framework, we built a high-quality Middle Eastern terrorist/extremist groups Web collection and benchmarked it against the US government Web site collection. The results showed that terrorist/extremist groups adopted similar levels of Web technologies as US government agencies. Moreover, terrorists/extremists had a strong emphasis on multimedia usage and their Web sites employed significantly more sophisticated multimedia technologies than government Web sites. We also found that terrorists/extremists seem to be as effective as the US government agencies in terms of supporting communications and interaction using Web technologies. More specifically, terrorists/extremists make heavy use of Web forums to facilitate their communication and coordination.

Our study provides insights for policy-makers to better apply counter-terrorism measures on the Web. Our results showed that Internet technologies, especially forums and chat rooms, have become a major means for terrorists/extremists to reach out to a broad audience. They have invested a significant amount of efforts and technical expertise into building their Web infrastructure. Security and law-enforcement experts should pay more attention to terrorists/extremists' online communication. We identified very high level of communicative activities in terrorist/extremist forums in our collection. Some documents in our collection were not readable using conventional applications. Some of these documents might contain hidden information from terrorists/extremists. Monitoring and deciphering such hidden messages could help intervene terrorist/extremist communication and prevent terrorism

attacks. Furthermore, we believe that the proposed Dark Web research methodology could also contribute to the terrorism research domains. The richness of the Dark Web contents calls for more studies being devote to this domain to help enrich our understanding of terrorists/extremists' Internet usage, online propaganda campaigns, and their psychological warfare strategies.

We have several future research directions to pursue. First, we plan to experiment with better data analysis methods and collaborate with more terrorism/extremism domain experts to better analyse and interpret our study results. For example, for the content richness comparisons, we would like to conduct a more detailed study to compare the richness of terrorist/extremist Web sites to government Web sites based on the percentage of each type of media in the overall contents. We also plan to conduct a cross-comparison which takes both the TS and WI attributes into consideration to gain more insight about the correlation of these attributes. Second, we plan to cooperate with Web technology experts to further improve the DAWS by incorporating additional attributes and adjusting the relevant weights. Third, we plan to expand the scope of our study by conducting a comparative analysis of terrorist/extremist groups' Web sites across different regions of the world. We also plan to conduct a time series analysis study on the Dark Web to analyse the evolution and diffusion of terrorist/extremist groups' Web presence. Last but not least, we also plan to explore more advanced machine-learning techniques to detect the technology and media usage pattern in terrorist/extremist Web sites to gain more insights into terrorists/extremists' technology usage.

Acknowledgments

This research has been supported in part by the following grants:

- NSF, "COPLINK Center: Information & Knowledge Management for Law Enforcement," July 2000–September 2005.

- NSF/ITR, “COPLINK Center for Intelligence and Security Informatics Research—A Crime Data Mining Approach to Developing Border Safe Research,” September 2003–August 2005.
- DHS/CNRI, “BorderSafe Initiative,” October 2003–March 2005.

We would like to thank Dr. Joshua Sinai formerly at the Department of Homeland Security, Al Qaeda expert Dr Marc Sageman, Dr. Chip Ellis from the Memorial Institute for the Prevention of Terrorism, and all the other anonymous domain experts for their insightful comments and suggestions on our project. We would also like to thank all members of the Artificial Intelligence Lab at the University of Arizona who have contributed to the project, in particular Homa Atabakhsh, Cathy Larson, Chun-Ju Tseng, and Shing Ka Wu.

References

- Anderson, A., 2003. Risk, terrorism, and the Internet. *Knowledge, Technology & Policy* 16 (2), 24–33 Summer 2003.
- Armstrong, H.L., Forde, P.J., 2003. Internet anonymity practices in computer crime. *Information Management & Computer Security* 11 (5), 209–215.
- Becker, A., 2004. Technology and Terror: the New Modus Operandi. *Frontline*, available at <http://www.pbs.org/wgbh/pages/frontline/shows/front/special/tech.html>
- Bowers, F., 2004. Terrorists spread their messages online. *Christian Science Monitor*, July 28, 2004, available at <http://www.csmonitor.com/2004/0728/p03s01-usgn.htm>.
- Bunts, G.R., 2003. *Islam in the Digital Age: E-Jihad, Online Fatwas and Cyber Islamic Environments*. Pluto Press, London.
- Chakrabarti, S., van den Berg, M., Dom, B., 1999. Focused crawling: a new approach to topic-specific Web resource discovery. In: *Proceedings of the 8th International World Wide Web Conference*, Toronto, Canada.
- Chen, H., Qin, J., Reid, E., Chung, W., Zhou, Y., Xi, W., Lai, G., Bonillas, A. A., Sageman, M., 2004. The Dark Web Portal: Collecting and Analyzing the Presence of Domestic and International Terrorist Groups on the Web. In: *Proceedings of International IEEE Conference on Intelligent Transportation Systems*.
- Chou, C., 2003. Interactivity and interactive functions in web-based learning systems: a technical framework for designers. *British Journal of Educational Technology* 34 (3), 265–279.
- Coll, S., Glasser, S.B., 2005. Terrorists Turn to the Web as Base of Operations. *Washington Post*, Aug 7, 2005.
- Demchak, C., Friis, C., La Porte, T.M., 2001. Webbing governance: national differences in constructing the face of public organizations. In: Garson, G.D. (Ed.), *Handbook of Public Information Systems*. Marcel Dekker, NYC.
- Denning, D.E., 2004. Information operations and terrorism. *Journal of Information Warfare* (draft), available at <http://www.jinfowar.com>.
- FBIS, 1995. Arab Afghans Said to Launch Worldwide Terrorist War. *Paris al-Watan al-Arabi*, December 1, 1995, pp. 22–24, FBIS-TOT-96-010-L.
- Hillman, D.C.A., Willis, D.J., Gunawardena, C.N., 1994. Learner-interface interaction in distance education: an extension of contemporary models and strategies for practitioners. *The American Journal of Distance Education* 8 (2), 30–42.
- Internet Haganah, 2005. Internet Haganah report, 2005, available at http://en.wikipedia.org/wiki/Internet_Haganah.
- ISTS, 2004. Examining the Cyber Capabilities of Islamic Terrorist Groups. Report, Institute for Security Technology Studies. <http://www.ists.dartmouth.edu/>
- Jackson, B.J., 2001. Technology acquisition by terrorist groups: threat assessment informed by lessons from private sector technology adoption. *Studies in Conflict & Terrorism* 24, 183–213.
- Jenkins, B.M., 2004. World Becomes the Hostage of Media-Savvy Terrorists: Commentary. *USA Today*, August 22, 2004. <http://www.rand.org/>
- Jesdanun, A., 2004. WWW: Terror’s Channel of Choice. *CBS News*, June 20, 2004.
- Kelley, J., 2001. Terror Groups Hide Behind Encryption. *USA Today*, Feb 5, 2001, available at <http://www.usatoday.com/tech/news/2001-02-05-binladen.htm>
- Muriel, D., 2004. Terror Moves to the Virtual World. *CNN News*, April 8, 2004, available at <http://edition.cnn.com/2004/TECH/04/08/internet.terror/>.
- Nunnally, J., 1978. *Psychometric Theory*. McGraw Hill, New York.
- Palmer, J.W., Griffith, D.A., 1998. An emerging model of Web Site design for marketing. *Communications of the ACM* 41 (3), 45–51.
- Preece, J., 2000. *Online communities: designing usability, supporting socialability*. Wiley, New York City.
- Reid, E., Qin, J., Chung, W., Xu, J., Zhou, Y., Schumaker, R., Sageman, M., Chen H., 2004. Terrorism Knowledge Discovery Project: a Knowledge Discovery Approach to Addressing the Threats of Terrorism. In: *Proceedings of Second Symposium on Intelligence and Security Informatics, ISI 2004*, Tucson, Arizona.
- Reilly, B., Tuchel, G., Simon, J., 2003. Political Communications Web Archiving: Addressing Typology and Timing for Selection, Preservation and Access. In: *Proceedings of the European Conference on Digital Libraries*.
- Schneider, S.M., Foot, K., Kimpton, M., Jones, G., 2003. Building thematic web collections: challenges and experiences from the September 11 Web Archive and the Election 2002 Web Archive. In: *Proceedings of the Third ECDL Workshop on Web Archives*, Trondheim, Norway, August 2003.
- Thomas, T.L., 2003. Al Qaeda and the Internet: The Danger of ‘Cyberplanning. *Parameters*, Spring 2003, pp. 112–123, available at <http://carlisle-www.army.mil/usawc/Parameters/03spring/thomas.htm>.
- Trevino, L.K., Lengel, R.H., Daft, R.L., 1987. Media symbolism, media richness, and media choice in organizations: a symbolic interactionist perspective. *Communication Research* 14 (5), 553–574.
- Tsfati, Y., Weimann, G., 2002. *www.terrorism.com: Terror on the Internet*. *Studies in Conflict & Terrorism* 25, 317–332.
- Weimann, G., 2004. *www.terror.net: How Modern Terrorism Use the Internet*. Special Report, US Institute of Peace, 2004, Available at <http://www.usip.org/pubs/specialreports/sr116.pdf>.
- Whine, M., 1999. *Cyberspace: A New Medium for Communication, Command and Control by Extremists*, available at <http://www.ict.org.il/articles/cyberspace.htm>
- Zhou, Y., Reid, E., Qin, J., Chen, H., Lai, G., 2005. US domestic extremist groups on the Web: link and content analysis. *IEEE Intelligent Systems (Special Issue on Homeland Security)* 20 (5), 44–51.