

The Dark Web Forum Portal

From Multi-lingual to Video

Hsinchun Chen¹, Dorothy Denning², Nancy Roberts², Catherine A. Larson¹, Ximing Yu¹, and Chun-Neng Huang¹

¹Management Information Systems Department, The University of Arizona, Tucson, AZ USA

²Department of Defense Analysis, Naval Postgraduate School, Monterey, California, USA USA

Corresponding author: cal@eller.arizona.edu

Abstract—Counter-terrorism, intelligence analysts, and other investigators continue to analyze the Internet presence of terrorists, hate groups, and other extremists through the study of primary sources including terrorists’ own websites, videos, chat sites, and Internet forums. Forums and videos are both particularly rich sources of information. Forums -- discussion sites supporting online conversations -- capture each conversation in a “thread” and the ensuing postings are usually time-stamped and attributable to a particular online poster (author). With careful analysis, they can reveal trends in topics and discussions, the sequencing of ideas, and the relationships between posters. Videos gain a global audience when posted to YouTube, but identifying and finding videos relating to a specific interest or topic can be difficult among the tens of millions of available items. The Dark Web Forum Portal was originally constructed to allow the examination, from a broad perspective, of the use of Web forums by terrorist and extremist groups. The Video Portal module has been added to facilitate the study of video as it is used by these groups. Both portals are available to researchers on a request basis. In this paper, we examine the evolution of the Dark Web Forum Portal’s system design, share the results of a user evaluation, and provide an overview of the development of the new video portal.

Keywords—counter-terrorism; social media; videos; text-mining; web-mining

I. INTRODUCTION

It has been recognized for several years that Web 2.0 social media have become a convenient platform for terrorists and extremists to exchange information and spread their ideas. Counter-terrorists, intelligence analysts, and other governmental, non-governmental, and academic investigators and researchers continue to analyze the Internet presence of terrorists, hate groups, and other extremists through the study of primary sources including terrorists’ own websites, videos, chat sites, and Internet forums. For example, the United Nation’s Counter-terrorism Implementation Task Force in 2009 issued a report describing member states’ concerns about continued terrorist use of the Internet for fundraising, recruitment, and cyber attacks, among other things, and analyzed steps to address this use [1]. McNamee et al. [2] examined the message themes found in hate group websites to

understand how these groups recruit and react to threats through the formation of group identity. Post [3] noted how terrorists had created a “virtual community of hatred” and wrote of the need to develop a psychology-based counter-terrorism program to, in part, inhibit potential participants from joining, reduce support for these groups, and undermine their activities.

The videos and forums of these groups are particularly rich sources of information. Videos gain a global audience when posted to YouTube, but for investigation or research, identifying and finding videos relating to a specific interest or topic can be difficult among the tens of millions of posted items. The Video Portal was designed to facilitate the study of video as it is used by these groups. Previous work on the portal focused on developing a method for identifying and collecting extremist videos [4], and examining an automatic, feature-based framework for classifying videos once they have been collected [5].

Forums are discussion sites supporting online conversations; each conversation is captured in a “thread” and the ensuing postings are usually time-stamped and attributable to a particular online poster. With careful analysis, they can reveal, for example, trends in topics and discussions, the sequencing of ideas, and relationships between posters. The Dark Web Forum Portal (DWFP) was originally constructed by the University of Arizona Artificial Intelligence Lab to allow the examination, from a broad perspective, of the use of Web forums by terrorist and extremist groups. The DWFP provides web-enabled access to 29 important jihadist and other extremist web forums and currently archives approximately 15 million messages. The forum is multi-lingual in scope, including forums in Arabic, English, French, German, and Russian, and provides search, browse, and analysis functions. Previous papers related to the DWFP have examined system functions as well as the collection and updating activities needed to keep it current [6; 7]. Abbasi and Chen [8] performed deeper analyses using data extracted from the portal to refine affect and sentiment analysis methods. More recently, the portal has been in use and examined by students and other researchers training in counter-terrorism and other

intelligence analysis domains.

In this paper, we examine the evolution of the Dark Web Forum Portal's system design, share the results of the user's case studies performed from an intelligence analysis perspective using the DWFP, and provide an overview of the development of the new video portal.

II. BACKGROUND

Intelligence analysis is "the process of evaluating and transforming raw data acquired covertly into descriptions, explanations, and judgments for policy consumers" [9]. This analysis might be applied to any of several recognized intelligence sources; commonly cited sources include, for example, Signals Intelligence (SIGINT), the information derived from an electronic signal; Imagery Intelligence (IMINT), the study of an image and its context; and Open Source Intelligence (OSINT), the gathering and analysis of information that is generally available to the public.

The Dark Web Project is largely concerned with "open source" information or intelligence. As a long-term research program that aims to computationally study and understand the international terrorism (Jihadist) phenomena, we have collected for several years, as comprehensively as possible, all types of web content generated by international terrorist and extremist groups including web sites, forums, blogs, social networking sites, videos, etc. In addition, we have developed various multilingual data mining, text mining, and other techniques to perform link analysis, content analysis and classification, web metrics (technical sophistication) analysis, sentiment analysis, authorship analysis, and video analysis in our research (see <http://ai.arizona.edu/research/terror>).

The approaches and methods developed in this project are for purposes of contributing to and advancing the field of Intelligence and Security Informatics (ISI). Such advances help related stakeholders to perform terrorism research and facilitate international security and peace. Some of these stakeholders include, and have included over the years, researchers in a variety of computer- and social sciences-related fields, government agencies concerned with the analysis of open-source intelligence, and educators concerned with the education and training of future analysts. Over the years, we have fielded requests for access to the portal from a variety of organizations and individuals including over 250 researchers from academia, officers in law enforcement, military and intelligence personnel, and other interested individuals from government and industry. Most have been provided access to version 1.0, the most stable prototype version of the Dark Web Forum Portal. Researchers from the Naval Postgraduate School, in partnership with the project, have been granted access to version 2.0 and are contributing to the development of version 2.5. More information about NPS'

role and use of the portal will be given in the next section, where we examine the development and evolution of the portal. We will follow that section with case studies by two defense analysis students who examined the portal in the context of its potential utility for intelligence analysis.

III. SYSTEM DEVELOPMENT AND EVOLUTION OF THE DARK WEB FORUM PORTAL

A. The Dark Web Forum Portal

1) Version 1.0

The Dark Web Forum Portal (DWFP), introduced in [10], maintains a collection of 29 online Jihadist forums, which currently contains 14,297,961 messages and 1,553,122 threads from 362,495 authors. Our incremental spiders visit all forums in the collection every two weeks, go through all the sub-forum and thread list pages and collect all new threads and messages that appeared in the latest two weeks. By post-processing the collected pages using our own parsers, we update our collection in the database. The details of incremental spidering process are covered in [10]. The web portal provides access to our data collection for users. Functionalities such as forum browsing, searching, social network visualization, and translation are supported. Google API was chosen for the translation function as being the most highly adopted and powerful translation tool (<http://code.google.com/apis/language/translate/overview.html>); it presently provides translations of 59 languages. While the version 1.0 of DWFP [10] does provide basic access to our Dark Web forum data collection, there are concerns from users regarding the usability of the portal. Users of version 1.0 have reported the following major problems:

Inconsistency in display: in the portal, users are able to browse forums by author, by thread, by time or by topic, in separated pages. However, in each page, different types of information are provided, making it difficult for users to consistently obtain precisely the type of information they need.

Lack of sophisticated search functionality: search functionalities, which are the most frequently used functions, are scattered across several different pages, making it difficult for users to figure out how to use them.

Slow translation: as most content in the portal is in a language other than English, quick and accurate translation is in great demand by users. The original translation utility in version 1.0 performs its translation function message by message, an initial design that eventually proved to be slow and cumbersome for users.

2) Version 2.0

To address these major issues, the portal was rebuilt and released as version 2.0; it adopted a new design for the user

interface based on previous feedback. With the new design, the system functionalities were grouped into three categories:

- **Forum browsing:** following the format of most online forums, forums may be browsed either by thread or by message. Additional filtering functionalities allow users to refine threads or messages to a certain time period or to those posted by certain authors.
- **Forum searching:** a more streamlined search functionality is presented to users. A sidebar is also provided on each search results page so that users can further refine and filter their search results.
- **Social network visualization:** this module remains unchanged from version 1.0.

Each of the functionalities listed above is described here in more detail.

a) Forum Browsing

Forum browsing is organized either by threads (shown in Figure 1) or by messages. Relevant information for each thread or message is displayed in the listing page, while the detailed content of the threads and messages is accessible to users with just a click.

In addition, the portal allows users to filter the returned thread or message list pages. For both the thread and message listing pages, the threads and messages, respectively, can be filtered by the start or end time and by author.

Thread ID	Thread Title	# Posts	# Authors	Start By	Last Post
190302	تغريب التيم جويل - استكشاف صيف كذا استكشاف الأرض	1	1	hetaq	hetaq
190308	تعمير استوديو يا اليوم صيفك المزمور 2019	1	1	كثير حق	كثير حق
190306	الامارة // التقرير الثاني السنوي لولاية حول الفساد - 20/12/2019	1	1	تد الرصي	تد الرصي
190325	الامارة // التقرير الاخباري عن الولاية 2019 - 21/12/2019	1	1	تد الرصي	تد الرصي
190304	الامارة // إعلان إسقاط طائرة تحبس بلا طيار في صحراء - 21/12/2019	1	1	تد الرصي	تد الرصي
190323	التحديات التي تواجه الخطبات الاستهلامية	1	1	صنطوق ابي	صنطوق ابي
190322	رسالة تد رملاني	2	2	جوياف	جوياف
190321	...عن المطالعة: شاهد (صين تطبات افروا انصروا الامم) صف الثانية لاصفيا 22/12	1	1	تد الرصي	تد الرصي
190320	صحة مطوق فوراياكو فري تمارب	1	1	تد الرصي	تد الرصي

Figure 1. Screenshots of forum browsing by threads

b) Forum Searching

To streamline the search functionality, both quick search and advanced search functionalities are available. For quick search, a search box is placed at the top of each page, allowing users to quickly search within the thread titles across all forums. The advanced search allows users to specify restrictions such as time, forum, and author, and to compose complicated queries using Boolean logic. Users may also choose to search within the contents of all messages. A

sample search result listing is shown in Figure 2. On the left, there is a sidebar grouping the search results by forum, by author and by year. These filters allow searches to be easily refined according to the user's evolving needs and preferences.

c) Translation

A new Google translation toolbar was adopted into this version of the portal. It quickly and efficiently translates the content of the entire web page into the language selected by the user.

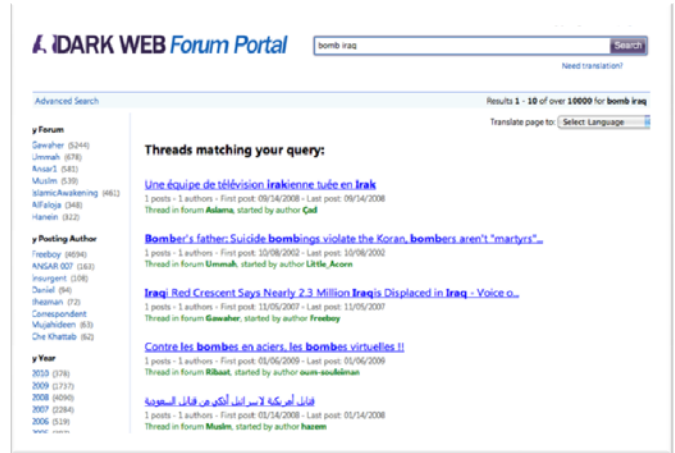


Figure 2. Results of the “quick search” bomb+iraq across all thread titles

3) Version 2.5

While version 2.0 addressed many issues in usability, improvements in searching were still needed. Search is one of the most important and well-used functions in the portal and as of version 2.0, the search results were still not very satisfactory in the following aspects:

- **Query parsing:** while version 2.0 added some Boolean searching capability, it did not support complex, sophisticated queries.
- **Search ranking:** the search ranking was problematic when multiple keywords with the “OR” relationship were entered by users.
- **Hit highlighting:** matched keywords were not always correctly highlighted; some highlighted words did not match the input search terms.

Given these issues, we embarked on a newer version of the portal, version 2.5, based on version 2.0. For this newer version, we adopted Lucene, a popular Java-based full-text indexing framework for the indexing and searching of thread titles and message contents.

Bringing in Lucene has resulted in both advantages and disadvantages. Through user testing, we can see that searching

has been improved through better query parsing, searching ranking and hit highlighting, but at the cost of maintaining and updating an index, separate from the database.

B. Conclusion

Figure 3 summarizes the evolution of the DWFP from version 1.0 to 2.5. By developing version 2.0 and 2.5 of the portal, we have made the DWFP a useful and powerful tool for users looking for topics and trends of interest in the Jihadist forums, thus further supporting its potential use as an open source intelligence tool. In a later section of this paper, we discuss the use of the DWFP by students learning to perform intelligence analysis tasks.

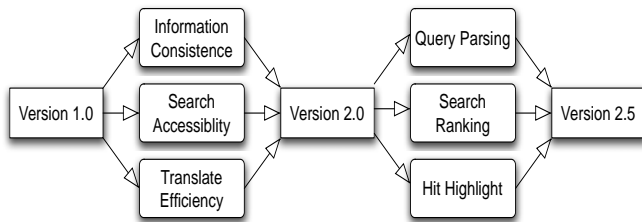


Figure 3. System evolution from version 1.0 through version 2.5.

IV. CASE STUDIES USING THE DARK WEB FORUM PORTAL

A. Introduction

Students taking a Defense Analysis course at the Naval Postgraduate School in Monterey, California developed the two case studies presented here. The students are generally mid-career military personnel, often with significant field or battle experience, from the United States or allied countries. They may be from any branch of the military and are seeking a graduate education through any of the 18 departments such as Defense Analysis, Information Sciences, etc. at the Naval Postgraduate School (NPS). The mission of NPS is to provide “high-quality, relevant and unique advanced education and research programs that increase the combat effectiveness of the Naval Services, other Armed Forces of the U.S. and our partners, to enhance our national security” (NPS website, <http://www.nps.edu/About/index.html>).

In its Defense Analysis course on Conflict and Cyberspace students examine how cyberspace, particularly the Internet, can serve as a tool, target, and source of conflict for both state and non-state actors. As part of the course, the instructor gave a demonstration of the Dark Web Forum Portal in class. The portal was presented as a potential resource for open source intelligence. Following the demonstration, the students were assigned a project to use the portal to investigate a topic of their choice, write a 1,200 word paper describing what they did and learned, and discuss their findings in a later class. The following two subsections discuss two of the student projects. In both cases, the students had significant previous military

and field experience and expected to continue their military service.

B. Case Study I: Dark Forums in Eastern Afghanistan: How to Influence the Haqqani Audience

In the first case study, “Dark Forums in Eastern Afghanistan: How to Influence the Haqqani Audience,” the student, who had had previous deployments to Afghanistan, postulated that one reason the war was continuing to drag on was that the U.S. was “losing the information war” with the people of Afghanistan and Taliban “safe haven” sites, an argument that has also been made by counter-terrorism researchers (e.g., [11]). The student used the Dark Web Forum Portal to investigate how the Haqqani Network of the Taliban, identified as one of the chief adversaries of the U.S., was able to twist U.S. tactical victories into operational defeats in the media and idea battlefields. After investigating relevant forums and postings, he cited, for example, patently false propaganda distributed through the forums which described American deaths that never actually occurred – victories claimed for the Taliban that never happened. The student also learned through forums on the portal that his own patrol had killed the son of a senior Haqqani member while on patrol. This was not information they had access to while in Afghanistan, and was extremely important tactical information that they would have been able to capitalize on while there. Other useful tactical information gleaned through postings on the portal included movements of other Haqqani members in the province, and previously unknown associations and linkages between certain individuals and organizations. According to the student, these are examples of information gaps that, had such intelligence reached them while in-country, would have allowed U.S. patrols to choose different courses of action.

C. Case Study II: Psychological Operations

In the second case study, a student familiar with military information support operations (MISO) used the Dark Web Forum Portal to investigate and compare the popularity and use of various media in the Muslim community. Such knowledge can help guide decision-making and communications strategies both in and outside a country of interest. The student examined forum threads and postings related to broadcast media (radio and television) and paper media (brochures, leaflets, and handbills), as all such communication methods may be useful for information operations. He also investigated postings relating to “propaganda,” which were often accompanied by very negative perceptions whether attributed to the U.S. or other Western governments, or to Middle Eastern governments. Much propaganda was attributed to “Zionist” conspiracy groups. Most messages concerning propaganda concerned the message content and the lack of legitimacy of the originator.

One interesting and useful finding was that radio was discussed in the forums more than any other media, with comments to the effect that radio, particularly unlicensed or “pirate” radio, was an effective means for reaching Muslims. According to the student, the continuing usefulness of radio outweighs that of other media, although forum participants also expressed interest in using print media such as brochures to spread their messages. Overall, the student thought that the DWFP could be a very useful tool for planning information operations in cyberspace as well as for maintaining situational awareness while in areas dominated by extremists.

D. Conclusion

In the U.S., the perceived value of open source intelligence has varied over time, peaking during the years of the Cold War and gaining in popularity again following 9/11. While OSINT cannot supplant information gained through other means, it can at times prove to be a useful supplement or corroboration.

V. VIDEO PORTAL MODULE

As described above, Web 2.0 social media have become a convenient platform for terrorists to exchange information and spread their ideas. Terrorist-related content obtained from the Internet is an important source of information for researchers trying to understand how terrorists think and how they act. Many previous studies focused on text-based social medias such as forums [12] and blogs [13]. Zhang and Zeng et al. [12] further provided a framework to efficiently collect postings of various terrorist related forums into a central repository to facilitate Dark Web research. However, as an important part of social media, online video-sharing websites have attracted less attention.

Videos are informationally dense and we believe they will become an important source to study for Dark Web researchers. There is little previous work in the area of dark web videos. Huang and Fu et al. [14] engaged in dark video identification and proposed a framework for classifying dark videos and segregating them out from non-dark videos. Their results implied that many dark videos existed on video-sharing websites without being immediately removed by the security mechanisms provided by the websites. However, these dark videos may eventually be removed by the websites before further studies are conducted. To facilitate research of dark videos, we proposed a framework, the Dark Video Portal, to collect relevant videos from video-sharing sites and develop a user-friendly interface which would allow users to search, query, and download the dark video collection.

1) System Design

The Dark Video Portal consists of three main components: data acquisition, data preparation, and system functionality.

Details of each component are shown in Figure 4 and described in the following sections.

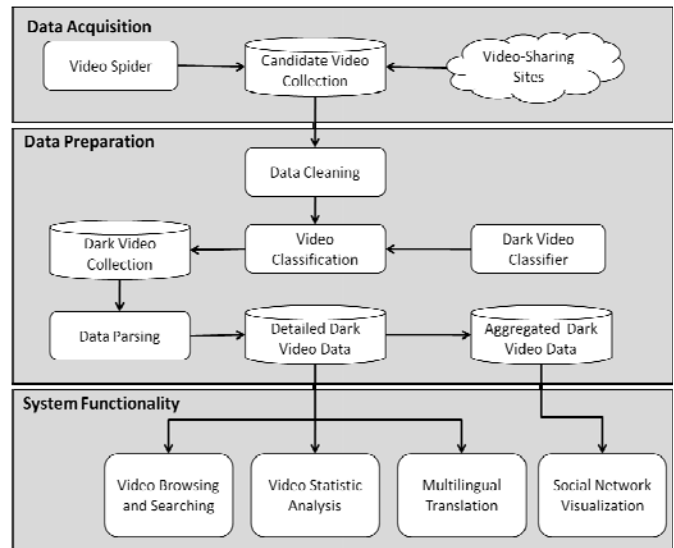


Figure 4. System Design of the Dark Video Portal

a) Data Acquisition

For data acquisition, spiders were developed to collect videos from the video-sharing sites. Different from dark forums where the majority of posts are related to terrorists, video-sharing sites are open to the public and all manner of videos are posted; dark videos form only a small portion of all available videos. Therefore, to efficiently identify Dark Videos, we follow the framework proposed by [14]. Spiders were developed to query video-sharing websites using sets of keywords predefined by Dark Web domain experts (see section on Dataset below). Each set of keywords represents a sub-collection of dark videos. Videos of query results (called query videos) and their metadata are first collected. Metadata refers to the text information related to a video and includes comments, descriptions, video titles, etc. Then videos related to query videos (called related videos) and videos published by the posters of the query video (poster videos) and their metadata are added to the collection. All these videos are put together as the Candidate Video Collection.

b) Data Preparation

The purpose of this step is to identify dark videos from within the Candidate Video Collection and extract detailed information about them for later usage. As described in [14], video classification is used to identify dark videos. However, before video classification is performed, video filtering is first used to filter out some videos and reduce the effort needed by video classification. In some video-sharing sites (e.g., YouTube), videos are categorized into predefined categories. Videos belonging to categories irrelevant to Dark videos (e.g., Music, Sports, and Entertainment) can first be removed.



Figure 5. Search Results Using Keyword “Bomb”

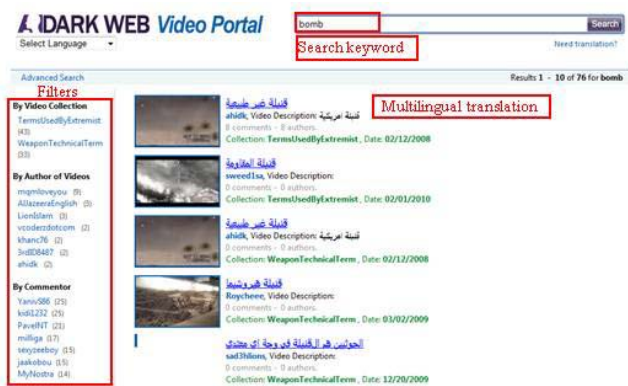


Figure 6. Detailed View of a Video

Video classification is then conducted to identify dark videos. A set of dark videos and ordinary videos are manually identified and used as a training dataset to build classification models. Training videos are represented by features extracted and selected from the metadata and fed into a support vector machine (SVM) to build a classifier. Then, following the same process, videos from the Candidate Video Collection are represented and classified by the classifier. Videos classified as dark videos then form the Dark Video Collection (see [14] for the details of video classification process).

Detailed information on each video in the Dark Video Collection is extracted and stored in the Detailed Dark Video Database. This information includes video names, video descriptions, name of the person posting the video, post dates, comments, commenters, and dates of comments. In addition, an Aggregated Dark Video Database is also constructed for the social network analysis by aggregating the detailed information of all videos.

c) Functionality

The Dark Video portal is implemented using Apache Tomcat and the database is implemented using Microsoft SQL Server 2000. Functionalities provided by the system include

video browsing and searching, video statistic analysis, multilingual translation, and social network visualization. In more detail, for statistic analysis, charts are created to show the trends of numbers of videos and comments posted over time. Google translate API (<http://code.google.com/apis/language/translate/overview.html>) is used to implement the translation function. The social network visualization function is realized by using JUNG (<http://jung.sourceforge.net/>) to show the interactions among video authors and commenters.

2) Dataset

Two sets of terrorist-related keywords are used to query video-sharing sites: terms used by extremists, and weapons technical terms. Terms used by extremists contains 22 English and 49 Arabic keywords, while the list of weapons technical terms consists of 2 English and 55 Arabic keywords. Table I shows keyword examples from each set.

TABLE I. KEYWORD EXAMPLES

Sub-collection	Keyword	Translation	Sub-collection	Keyword	Translation
Terms used by extremists	Islamic Jihad	-----	Weapons technical terms	تحضير متفجر	Explosive preparation
	Hamas	-----		موسوعة الأسلحة الكبرى	Great weapons encyclopedia
	Usbat Al Ansar	-----		الحزام الناسف	Explosives belt
	المجاهدين	Mujahideen			

As per Table II, 104,206 videos were collected by using these two sets of keywords to create the Candidate Dark Video Collection. After video classification, 12,728 videos (12%) were identified as dark videos and formed the Dark Video Collection. Table II shows the details of the Dark Video Collection.

TABLE II. KEYWORD EXAMPLES

Sub-collection	# of Candidate Videos	# of Dark Videos	# of Video Authors	# of Comments	# of Commenters
Terms used by extremists	90,725	11,355	4,743	420,446	172,146
Weapons technical term	13,481	1,373	767	42,210	22,061
Totals	104,206	12,728	5,530	462,656	194,207

3) Functionality

Four main functionalities are provided by the Dark Video portal. This section gives the details of each function.

a) Video Browsing and Searching

This system provides functionalities for users to browse and search videos within a sub-collection or cross sub-collections. Figure 5 shows the search result using the keyword “bomb.” In the example, 76 videos containing the term “bomb” in the video titles and descriptions have been returned. Search keywords are automatically translated into Arabic and English, so that videos with Arabic titles can also be searched by using English search terms. In addition, filter functions are also

provided. Users are allowed to refine their search results by sub-collections, video authors, commenters, or year(s). By clicking on a video title, details of the video are displayed as shown in Figure 6.

b) Video Statistics Analysis

Statistics and other general information pertaining to the numbers of videos posted over time, comments, and top video posters and commenters are provided for each sub-collection. Flash-based charts are created to show information across different time periods. In addition, users are allowed to aggregate information with different time windows (e.g., 1-day, 5-day, 1-month, and 6-months).

c) Multilingual Translation

Since the majority of the Dark Video Collection videos are in Arabic, a multilingual translation function similar to that implemented in the main Dark Web Forum Portal was integrated to translate web pages into different languages.

d) Social Network Visualization

Social network analysis can be used to understand interactions between video posters and commenters under a specific topic (e.g., Bomb). Following [12], we modeled the social network of videos by extending the definition of the reply network in forums. More specifically, every node represents a user (a video poster/commenter) and every link represents a comment relationship. If B makes a comment on a video posted by A, a directed link will be added into the social network, pointing from A to B.

Here we show an example of social network analysis which was to find the most active users who commented on Jihad-related videos between Oct. 2009 and Jan. 2010.

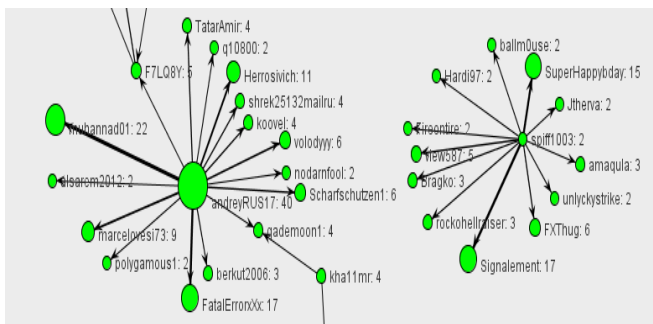


Figure 7. Social Network Analysis

First, the keyword “Jihad” was used to search for videos and comments related to this topic. As a result, 1,700 videos and 176,730 comments were identified; the top 5 commenters were Ialkimawi (62 comments), andreyRUS17 (40 comments), muhannad01 (22 comments), FatalErrorXx (17 comments), and Signalement (17 comments). After applying a filter which showed only the top 20% of users by using

number of comments, Ialkimawi has no connections anymore, indicating that this user only comments on users who seldom post comments or videos. Figure 7 shows the filtered results of social network analysis. The results show that the cluster on the left contains andreyRUS17, muhannad01 and FatalErrorXx, while the cluster on the right contains Signalement. In addition, the thickness of the link indicates that muhannad 01 and FatalErrorXx interacted a great deal with andreyRUS17, while Signalement and SuperHappybday interacted quite a bit with spiff1003.

e) Conclusion

The Video Portal has only recently been released to researchers and is intended to serve as another potential resource for open source intelligence. Given the difficulties of identifying and locating these videos, particularly since they are routinely censored from video sharing sites, and the great deal of interest that has been expressed by various groups in accessing them, we believe our intended audience members of researchers, analysts, and other investigators will find the video module of the portal to be useful. A period of evaluation will shortly ensue to help us gauge interest on the part of the community as well as usability.

REFERENCES

- [1] United Nations, Counter-Terrorism Implementation Task Force (CTITF), Report of the Working Group on Countering the Use of the Internet for Terrorist Purposes, February 2009. (http://www.un.org/terrorism/pdfs/wg6-internet_rev1.pdf)
- [2] McNamee, L. G., Peterson, B. L., and Peña, J. "A call to educate, participate, invoke and indict: understanding the communication of online hate groups," *Communication Monographs* 77: 2, pp. 257-280, 2010.
- [3] Post, J. "'When hatred is bred in the bone:' the social psychology of terrorism," *The Annals of the New York Academy of Sciences*, 1208, Psychiatric and Neurologic Aspects of War, pp. 15-23.
- [4] Fu, T. J., Huang, C., and Chen, H., "Identification of extremist videos in online video sharing sites," in *Proceedings of the IEEE International Intelligence and Security Informatics Conference (Dallas, Texas, June 8-11, 2009)*.
- [5] Huang, C., Fu, T. J. and Chen, H., "Text-based video content classification for online video-sharing sites," *Journal of the American Society for Information Science and Technology*, 61:5, pp. 891-906, 2010.
- [6] Zhang, Y., Zeng, S., Fan, L., Dang, Y., Larson, C. and Chen, H. "Dark Web Forum Portal: searching and analyzing jihadist forums," in *Proceedings of the IEEE International Intelligence and Security Informatics Conference (Dallas, Texas, June 8-11, 2009)*.
- [7] Fu, T.J., Abbasi, A. and Chen, H., "A focused crawler for Dark Web forums," *Journal of the American Society for Information Science and Technology*, 61:6, pp. 1213-1231, 2010.
- [8] Abbasi, A., and Chen, H. "CyberGate: a system and design framework for text analysis of computer mediated communication," *MIS Quarterly (MISQ)*, 32: 4 (Special Issue on Design Science Research), pgs. 811-837, 2008.
- [9] Lefebvre, S. "A look at intelligence analysis." *International Journal of Intelligence and Counterintelligence*, 17:2, pp. 231-264, 2004.

- [10] Zhang, Y. and Zeng, S. et al., "Developing a Dark Web collection and infrastructure for computational and social sciences," *Intelligence and Security Informatics (ISI)*, 2010 IEEE International Conference on Vancouver, BC, Canada, pp. 59 - 64, 2010.
 - [11] Betz, D., "The virtual dimension of contemporary insurgency and counterinsurgency," *Small Wars and Insurgencies*, 19:4, pp. 510-540, 2008.
 - [12] Zhang, Y., and Zeng, S. et al., "Developing a Dark Web collection and infrastructure for computational and social sciences," *Intelligence and Security Informatics (ISI)*, 2010 IEEE International Conference on Vancouver, BC, Canada, pp. 59 - 64, 2010.
 - [13] Yang, C.C. and Ng, T.D., "Terrorism and Crime Related Weblog Social Network: Link, Content Analysis and Information Visualization," *Intelligence and Security Informatics (ISI)*, IEEE International Conference on Intelligence and Security Informatics, NJ, pp. 55-, 2007.
 - [14] Huang., C. N., Fu, T., and Chen, H., "Text-based Video Content Classification for Online Video-sharing Sites," *Journal of the American Society for Information Science and Technology (JASIST)*, 61:5, pp. 891-906, 2010.
-