Chapter 12

Analyzing Deviant Socio-Technical Behaviors Using Social Network Analysis and Cyber Forensics-Based Methodologies

Samer Al-khateeb, Muhammad Hussain, and Nitin Agarwal

Contents

12.1 Introduction 264
12.2 Literature Review 265
12.3 Methodology 266
12.4 Case Studies 269
12.4.1 DAESH or ISIS/ISIL Case Study: Motivation, Evolution, and Findings 269
12.4.1.1 Exploring the Network of the Top Disseminators of ISIL 269
12.4.1.2 Beheading of Innocent People by ISIL 272
12.4.2 Novorossiya Case Study: Motivation, Evolution, and Findings 274
12.5 Conclusion and Future Work 279
Acknowledgments 279
References 279

Online social networks (OSNs) have grown exponentially in a short period of time, and this growth has revolutionized how societies interact with each other. Many people around the world use social media on a daily basis; for example, there are about 1.3 billion registered Twitter users, with an average of 100 million daily active users, and 65 million users just in the United States [1]. In addition to Twitter, Facebook is the largest social network in the world. The social networking site comprises approximately 1.65 billion monthly active users with about 167 million daily active users in the United States and Canada who spend an average of 20 minutes of their day on Facebook [2].

12.1 Introduction

The use of social media, specifically, went from being a source of entertainment or a way to find and connect with friends or family members, no matter where they are globally, to more of a deviant usage/purpose, e.g., to conduct cybercrime, hacking, cyber terrorism, spread propaganda or misinformation, conduct cyber warfare tactics, or other similar deviant acts. For example, the Russians are spending millions of dollars to finance the Kremlin's Troll Army (legions of pro-Russia, English-speaking, Internet commenters) to promote President Vladimir Putin and his policies, and to spread disinformation about some events or disseminate a propaganda war on the Ukraine [3]. Often these acts exhibit a flash mob style behavior, i.e., a process where a group of individuals get together in cyberspace, conduct a deviant behavior, and then disappear into the anonymity of the Internet. We call such behavior deviant cyber flash mobs (DCFM) [4].

These nefarious uses of social media pose a significant threat to society, and thus require special research attention. It would be of benefit to the information assurance domain, and its respective subdomains, to conduct new research initiatives into the phenomenon of deviant behavior in OSNs especially with the vast amounts of evidentiary information that is continuously generated on different social media outlets. In this chapter, we study the following research questions:

- What strategies and tools do deviant groups, e.g., transnational crime organizations or terrorist groups, use to disseminate their propaganda? And who is responsible for disseminating it (e.g., powerful actors)?
- Can we measure an individual's interest, control, and power in the dissemination process using the theoretical constructs of collective action, thereby modeling individual motivations to participate in propaganda dissemination?
- Are botnets involved in the dissemination process and how sophisticated are these bot networks? What role do these bots play in such information maneuvers? Are there structural pattern(s) among bot networks? How can we detect them?

Seeking answers to the aforementioned research questions, we make the following contributions in this chapter:

- We develop a systematic methodology that can be followed to analyze propaganda dissemination. This methodology was obtained from several experiments we conducted on the dataset we collected for different events mentioned above.
- We identified the strategies and tools that are used by deviant groups to conduct such a deviant act, e.g., propaganda dissemination.
- We show how cyber forensics can be used to discover hidden connections between information actors and can be used to study the cross media affiliations.

The rest of the chapter is organized as follows. We provide a brief literature review in Section 12.2. Section 12.3 discusses our methodology. We discuss the two cases we have investigated (namely, Daesh or ISIS/ISIL and Novorossiya) along with the results and analysis we obtained in each case in Section 12.4. Finally, Section 12.5 summarizes the study with possible future research directions.

12.2 Literature Review

In our methodology we applied an algorithm called focal structure (FSA) which is an algorithm that was developed by Sen et al. [5] to discover a set of influential nodes in a large network. This set of nodes does not have to be strongly connected and may not be the most influential on its own but by acting together it forms a compelling power. This algorithm was tested on many real world information campaigns such as the Saudi Arabian women's *Oct26Driving* campaign on Twitter* and during the 2014 Ukraine Crisis† when President Viktor Yanukovych refused to sign a European association agreement.

Botnets/bots/or automated social actors/agents (ASAs) are not a new invention. They have been used since 1993 in the Internet relay chat (IRC), and known as Eggdrop. They used to do very simple tasks such as greeting new participants and warning them about the other users' actions [6]. Then the usage of botnets evolved over time due to their multi-functionality that can be performed and their easiness to implement. In our work we were able to identify and study the network structure (the way the network looks) of botnets in all of the aforementioned events. A similar study was conducted on the Syrian Social Bot (SSB) that was used to disseminate propaganda during the Syrian civil war in 2012 [7]. Abokhodair et al. have categorized the bots by their activity type into four categories, namely: Core

* The right to drive campaign #oct26driving (available at: http://bit.ly/1OmyCIO).
† Ukraine protests after Yanukovych EU deal rejection (available at: http://bbc.in/1qhcy6V).

bots—Generators (tweet a lot), Core bots—Short Lived (retweet a lot and were active for less than six weeks), Core bots—Long Lived (retweet a lot and were active for more than 25 weeks), and Peripheral Bots (retweeting one or more tweets generated by the core bots and the account names look like a legitimate user account name).

We have also used some cyber forensics tools and techniques to discover the hidden relationships between different blog sites. We used Maltego tool, which is an open source intelligence and forensics application. It saves a lot of time in mining and gathering of information as well as the representation of this information in an easy to understand format. In addition to that we used some cyber forensics techniques such as Google Analytics ID, which is an online analytics tool that allows a website owner to gather some statistics about their website visitors such as their browser, operating system, and country. Multiple sites can be managed under a single Google analytics account. The account has a unique identifying "UA" number, which is usually embedded in the website's code [8]. Using this code other blog sites that are managed under the same UA number can be identified. This method was reported in 2011 by *Wired*, and also was cited in the book *Open Source Intelligence Techniques* by Michael Bazzell, an FBI cyber crime expert [8,9].

12.3 Methodology

In this section, we present our methodology that we followed to obtain the results and findings mentioned in Sections 12.4.1 and 12.4.2. Figure 12.1 shows our methodology as a flowchart of operations we used in combination with the software. This is followed by a stepwise explanation of each step/component in the diagram. We used the following software in our methodology:

- **Maltego:** a cyber forensics tool that helps uncover the hidden connection between different blogs sites, available at: http://bit.ly/1OoxDCD.
- **GoogleTAGs:** for collecting data in a continuous manner, available at: http://bit.ly/1KPrRH2.
- **TAGSExplorer:** to have a live visualization of the data collected with GoogleTAGs, available at: http://bit.ly/24NmFjy.
- **Blogtrackers:** to analyze the blogs whose data we collected. The tool can be accessed via the following URL: blogtrackers.host.ualr.edu.
- **CytoScape:** an open source software platform for data visualization, available at: http://bit.ly/1VTWOow.
- **NodeXl:** to collect and analyze the data, available at: http://bit.ly/1WKA5u9.
- **Linguistic Inquiry and Word Count (LIWC):** to calculate the sentiments scores, available at: http://bit.ly/1WKAyN3.
- **IBM Watson Analytics:** to explore the dataset and get further insights, such as the nature and type of conversations, available at: http://ibm.co/214CjoD.

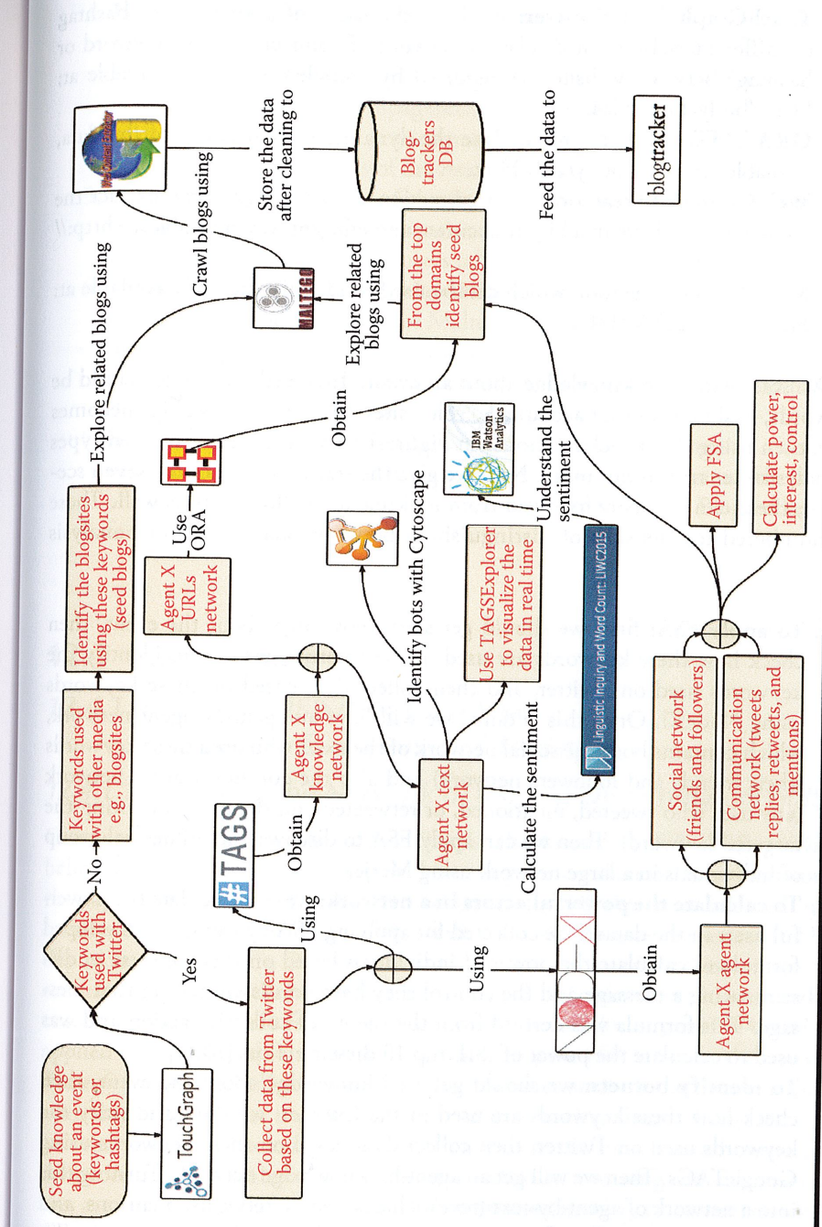


Figure 12.1 Flowchart of the methodology we followed to study the two cases.

- **TouchGraph SEO Browser:** to check the usage of a keyword or hashtag on different websites, it displays a network of connectivity (by keyword or hashtag) between websites, as reported by Google's database, available at: http://bit.ly/1Tm8Cz4.
- **ORA-LITE:** to assess and analyze the dynamic meta-network of the data, available at: http://bit.ly/27fuHnv.
- **Web Content Extractor:** a web extraction software, which can extract the content of a website in a highly accurate and efficient way, available at: http://bit.ly/1uUtpeS.
- **Merjek:** an online tool, which can be used to identify the FSA, available at: http://bit.ly/21YV5OC.

We start with seed knowledge about an event. This seed knowledge could be keywords, hashtags, twitter accounts, or blog sites. Once this knowledge becomes clear, then using the flowchart shown in Figure 12.1 we can perform seven types of analysis/scenarios if not more. Next we give the steps for each of the seven scenarios; these can be easily followed from looking at the flowchart as well. These are numbered for the sake of distinguishing between each scenario or analysis type:

1. **To apply FSA:** first, we should get seed knowledge about the event, then check how these keywords are used in the Internet; we should identify the keywords used on Twitter, and then collect data based on these keywords using NodeXl. Once this is done, we will have an agent-by-agent network, which contains both the social network of the users who used these keywords (their friends and followers network) and also the communication network (any user who tweeted, mentioned, or retweeted and the text containing the targeted keyword). Then we can apply FSA to discover the influential group of individuals in a large network using Merjek.
2. **To calculate the powerful actors in a network:** we can calculate the powerful users in the dataset we collected for applying FSA by using our developed formula to calculate the power of individuals based on their interest in disseminating a message and the control they have on disseminating that message. This formula was derived from the theory of collective action and was used to calculate the power of ISIL top 10 disseminators [10,11].
3. **To identify botnets:** we should get seed knowledge about the event, then check how these keywords are used in the Internet; we should identify the keywords used on Twitter, then collect data based on these keywords using GoogleTAGs. Then we will get an agent-by-knowledge network. Turning this into a network of agent-by-text (text includes tweets, retweets, mentions, and replies) and using CytoScape we can identify botnets in the network. We used this method to identify botnets that were working to disseminate ISIL beheading video propaganda [4].

4. **Understanding the sentiments:** we can use the same network mentioned in the previous step to understand the sentiments of the public about a specific event, e.g., how does the public feel about an issue or what kind of narrative is being pushed in the propaganda? This can be done by using LIWC to calculate the score of the sentiments of the text, and then use IBM Watson Analytics to ask questions of interest, and then find an answer to it.
5. **Have a live visualization of the data:** on the same data collected in Step 3 we can use TAGSExplorer to continuously collect data and have a live visualization of the Twitter feed of users talking about the event.
6. **Using Maltego to discover related blogs sites:** anytime we have a seed blog or a Twitter handle we can use Maltego to discover the other blogs that are owned by the same person or managed by the same unique identifying "UA" number. We can also find blogs from the Twitter handle or vice versa to study the cross media affiliation using Maltego.
7. **Blog analysis using Blogtrackers:** once we identify the blogs of interest we can crawl their data using a web content extractor tool to extract the content of the blog's site and then clean the data and feed it to the Blogtrackers tool for further analysis.

12.4 Case Studies

12.4.1 DAESH or ISIS/ISIL Case Study: Motivation, Evolution, and Findings

12.4.1.1 Exploring the Network of the Top Disseminators of ISIL

Our interest in studying the Islamic State's (also known as ISIL, ISIS, or Daesh) behavior on social media started with a study we conducted on the network of the top 10 disseminators of ISIL on social media [11] that were released by the International Center for Study of Radicalization and Political Violence (ICSR) in September 2014 [12]. In Carter et al.'s study, they interviewed the disseminators and recruiters of ISIL on social media and published their identities. In our work, we crawled those recruiter's friends' and followers' network (in August 2014) and then applied our developed framework to identify the powerful actors in that network. *Powerful actors* are individuals who assert a lot of control on the dissemination of a message and possess a great interest in disseminating them [10,11]. We found out that the top disseminator nodes are not only the most central nodes (most connected) but they also constitute a focal structure, meaning the top disseminators are coordinating with each other. They are coordinating in disseminating ISIS propaganda, forming a powerful campaign organization network. Figure 12.2(a) and (b) show two tweets as samples of the tweets we collected from the ISIL network containing propaganda as well as a message to the their followers. We did a

Figure 12.2 This figure shows a sample of the tweets we collected from the Daesh/ISIL network containing propaganda and messages to their followers. (a) Propaganda tweeted by one of the top 10 disseminators we have in our datasets. (b) Tweet containing a message to the followers of that account.

follow-up on the accounts we collected in their network and found out that some of these accounts were suspended by Twitter, others were attacked by Anonymous* (a hacking group), while the owners of other accounts were physically captured by law enforcement. Therefore, we decided to recrawl the network on February 2015 (six months later) to find out how the suspensions and cyber attacks affected the overall network. We found that ISIL increased its recruiting activities to almost double. Also, we found out some accounts were mostly used/owned by more than one individual, e.g., the owner of *@shamiwitness* was captured by law enforcement but the account was still active and attracting more people as we recrawled the network. Figure 12.3 shows the old and new network after recrawling it by six months, where red nodes represent the nodes that still exist, white nodes are the nodes suspended by Twitter, and blue nodes are the new ones added. The big 10 nodes represent the top 10 disseminators of ISIL as identified by ICSR.

Since very few users share their location on Twitter, we tried to infer the location of those Twitter accounts using the time zone they follow which is available for most users we collected (see Figure 12.4). Figure 12.4 shows the top 10 disseminators marked in orange while the rest of their friends and followers are marked in green. We zoomed-in on two of the top 10 disseminators *@Shamiwitness* bottom left and *@Ash_shawqi* bottom right. We found out that many users do not follow the time zone where they actually live in which indicates that they are either using proxy servers to hide their physical location or they are responsible for disseminating messages to audiences in different time zones than the one they follow. For example, *@Shamiwitness* follows Central Africa time zone but he was captured in

* Anonymous "Hacktivists" strike a blow against ISIS (Available at: http://bit.ly/1vzrTSQ).

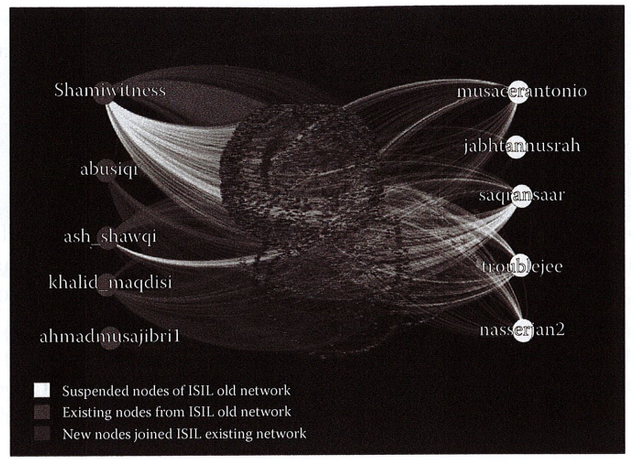


Figure 12.3 A total of 16,539 nodes and 21,878 edges (the friends/followers of ISIL's top 10 disseminators/powerful actors.

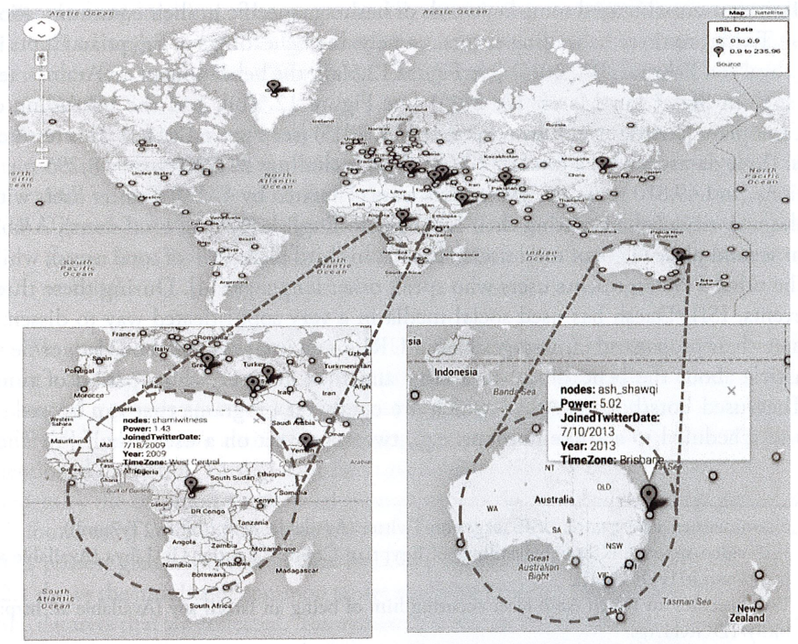


Figure 12.4 Inferring geolocation of ISIL's top 10 disseminators.

Bangalore, India.* We also found out that some accounts changed their identity after we recrawled the network, e.g., the account *@Ash_shawqi*: On the time of data collection this account had 5,990 tweets, 3,867 followers, and follows 279 users. His time zone is Brisbane, Australia. Six months later this account looks totally different. First, he wiped his old tweets and friends/followers (has just three tweets and one follower). Second, his profile language changed from Arabic to Russian. Third, he shared his location as in "Москва" which means "Moscow" in English. Finally, he used to describe himself as a member of Ahlu-Sunnah Wal Jama'ah (Platform of Tawhid and Jihad) but six months later his description changed into Russian language "Помощь должна совершаться не против воли того, кому помогает," which is translated using the Google translation service to English as "Aid should not be committed against the will of the one who helps."

12.4.1.2 Beheading of Innocent People by ISIL

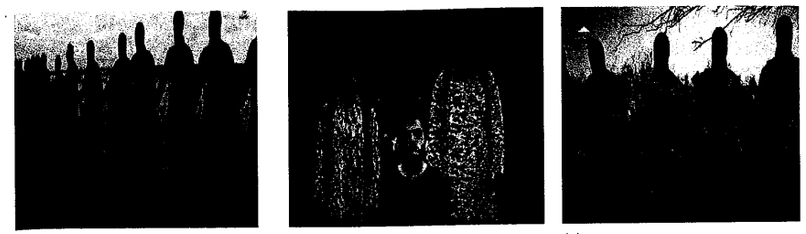
In 2015, the so-called Islamic State or Daesh started releasing videos on social media where they depicted gruesome beheadings of innocent people and accusing them of being unbelievers or traitors in an attempt to spread horror among the minority religious in the Middle East and capture the attention of the media. We crawled the communication network (tweets, retweets, replies, and mentions) of the Twitter users who used some keywords or hashtags specific in their communication on Twitter to three beheading events, namely: the beheadings of Egyptian Copts in Libya† (on February 15, 2015) [see Figure 12.5(a)], the beheading of an Arab-Israeli "spy" in Syria‡ (on March 10, 2015) [see Figure 12.5(b)], and the beheading of Ethiopian Christians in Libya§ (on April 19, 2015) [see Figure 12.5(c)]. This resulted in three datasets with a total of 80,685 texts including: 22,580 tweets, 8,295 mentions, and 49,810 retweets. These texts were generated by 47,148 Twitter users with the majority of them setting their language to English (67% of total users), Arabic in second place (17% of total users), French in third place (5% of total users), while the other 11% represents users who speak other languages [4]. During these three events, ISIL/Daesh has used social media in a very sophisticated way to disseminate their propaganda (messages have a URL to an image/YouTube video, or news article about the beheadings) to a large audience in a very short period of time. They used bots/botnets/ASAs, which are computer programs that can be tasked and scheduled to act like humans, e.g., tweet, retweet on a user behalf [4]. They

* Shami witness arrest rattles ISIS' cages on Twitter (Available: http://bit.ly/1Ty7um0).

† ISIS video appears to show beheadings of Egyptian Coptic Christians in Libya (Available at: http://cnn.it/1vO9CkA).

‡ ISIL executes an Israeli Arab after accusing him of being an Israeli spy (Available at: http://bit.ly/1DGRHAg).

§ Isis video purports to show massacre of two groups of Ethiopian Christians (Available at: http://bit.ly/1yJX8fp).



(a) (b) (c)

Figure 12.5 This figure shows a screenshot for each of the beheadings of innocent people executed by ISIL/Daesh. (a) The beheading of Egyptian Copts in Libya by ISIL on February 15, 2015. (b) The beheading of the Arab-Israeli "spy" in Syria by ISIL on March 10, 2015. (c) The beheading of the Ethiopian Christians in Libya by ISIL on April 19, 2015.

have used some of the very effective information maneuvers to disseminate their messages such as:

- Misdirection, where the bot tweets unrelated news that is happening somewhere else but mentions a hashtag related to the beheading crises.
- Hashtag-latching, when strategically associating unrelated but popular or trending hashtags to target a broader, or in some cases a very specific, audience (e.g., using the #WorldCup and then including a URL of a beheading video).
- Smoke screening, when the bot would mention something about ISIL but not necessarily related to the beheading (similar techniques have been used in the SSB to raise awareness of the Syrian civil war [7]).
- Thread-jacking, the change of topic in a "thread" of discussion in an open forum (e.g., using a hashtag of #ISIL but the tweet has a link to a shopping website).

We also observed that once the news released about the event many people talk about it and try to disseminate it to their friends/followers but two days later the activity started to decline (people know about the event already either by social media or some other means). We also intersected the datasets we collected about ISIL (the one mentioned in Section 12.4.1.1 and the three datasets mentioned in Section 12.4.1.2) to identify the common nodes in an attempt to find the new disseminators' network. We found out a lot of common nodes between datasets, which means those common users possess greater interest in disseminating the propaganda of ISIL/Daesh. Table 12.1 shows the number of common nodes resulting from the intersection ($\cap$) between ISIL datasets that we collected. The symbol ($\cap$) we use in the table is the mathematical intersection of two sets, i.e., the intersection between two sets, set "X" and set "Y" is the set "Z" which contains the unique elements from both set "X" and set "Y" [13].

Table 12.1 The Number of Common Nodes between the Datasets We Collected

Dataset Names	*# of Nodes*
Beheading of Coptic's ∩ beheading of Arab Israeli	265
Beheading of Coptic's ∩ beheading of Ethiopians	753
Beheading of Ethiopians ∩ beheading of Arab Israeli	339
Dataset of top 10 disseminators ∩ beheading of Arab Israeli	126
Dataset of top 10 disseminators ∩ beheading of Ethiopians	83
Dataset of top 10 disseminators ∩ beheading of Coptic's	61
Beheading of Coptic's ∩ beheading of Arab Israeli ∩ beheading of Ethiopians	68

Note: These common nodes have a great interest in disseminating ISIL's propaganda and might be the new disseminators of ISIL propaganda.

12.4.2 Novorossiya Case Study: Motivation, Evolution, and Findings

In this case study we studied the influence operations in Novorossiya. We followed a similar approach to that we followed in the ISIL case study, but here we added the cyber forensics aspect to reveal hidden connections between different organizing groups and to study the cross media affiliation of individuals and groups.

We started by studying the case of the Crimean water crisis* when the Russians invaded the Crimean peninsula on March 16, 2014. This crisis was met with international discontent and a sense of Russian imperialism to expand their reign of power. The United Nations as well as the NATO Secretary General condemned this expansion of Russian sphere of influence. Pro-Russian media propagandized that the closing of a main irrigation canal in April by Ukrainian authorities caused starvation in the peninsula and the death of many crops like rice, corn, and wheat. The pro-Russian media further emphasized that this has caused a humanitarian crisis, which was followed by grievances and requests for help. Many on-the-ground reports on developing conflicts and problems are reported in a variety of open source platforms including blogs, websites, Twitter, Facebook, and other open source channels such as YouTube. During this crisis, bots were used to effectively disseminate thousands of messages in relation to the Crimean water crisis. We collected data for the period between April 29, 2014 8:40:32 PM and July 21, 2014 10:40:06 PM UTC from Twitter using keywords related to this crisis and we wanted to investigate the tactical information maneuvers, especially the role of botnets in propaganda dissemination

* Aid Elusive, Crimea farms face hurdles (available at: http://nyti.ms/1UOpzlg).

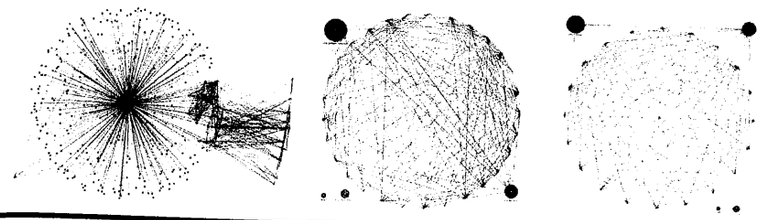


Figure 12.6 Naive botnets observed during the Crimean water crisis (2014). Mutual reciprocity and extremely closely knit behaviors were observed.

campaigns. This resulted in collection of 1,361 unique tweets, 588 unique Twitter users, and 118,601 relations/edges between the Twitter users. There are four basic types of relations in the Twitter data, namely, follows, mentions, replies, and tweets. We found out that these bots had a central account that is responsible for giving the propaganda to them and they work to disseminate these propaganda messages. By closely examining their network, as depicted in Figure 12.6, we found out that they have a mutually reciprocated relationship, suggesting the principles of "Follow Me and I Follow You" (FMIFY) and "I Follow You, Follow Me" (IFYFM) in practice—a well-known practice by Twitter spammers for link farming or quickly gaining followers [14–16]. The bots in this case were considered simple bots, which means if you find one you find others. While in the following case explained next we found more sophisticated bots, which were harder to discover.

After the study of the Crimean water crisis, we studied the propaganda projected against two military exercises conducted by U.S. forces and NATO, namely the Dragoon Ride Exercise* and the Trident Juncture Exercise† (TRJE 2015). In the Dragoon Ride Exercise, a march of U.S. soldiers were sent on a mission as part of Operation Atlantic Resolve and began Operation Dragoon Ride (in March 21, 2015) to exercise the unit's maintenance and leadership capabilities and demonstrate the freedom of movement that exists within NATO. That march covered more than 1,100 miles and across five international borders including Estonia, Latvia, Lithuania, Poland, and the Czech Republic. The Trident Juncture Exercise, which involved 36,000 personnel from more than 30 nations, took place throughout Belgium, Germany, the Netherlands, Norway, Spain, Portugal, Italy, Canada, the Mediterranean Sea, and the Atlantic Ocean to demonstrate NATO's capability and capacity to meet the present and future security challenges.

There was a lot of propaganda projected during these two military exercises. In addition to that, some of the local residents from the participating countries did not like those exercises. There were calls for civil unrest and a lot of propaganda asking

* Operation Atlantic Resolve exercises begin in Eastern Europe (available at: http://1.usa.gov/1rDSxcb, Last accessed: June 12, 2016).

† Trident Juncture 2015 (available at: http://bit.ly/1Q...

people to protest and conduct violent acts against the troops participating in both exercises. For the Dragoon Ride Exercise, this was done mainly by a group of botnets. These botnets were identified using Scraawl (an online social media analysis tool available at www.scraawl.com). We collected the network of these bots in the period between May 8, 2015 8:09:02 PM and June 3, 2015 11:27:31 PM UTC of 73 Twitter accounts that included friend–follower relations and tweet–mention–reply relations. This resulted in 24,446 unique nodes and 31,352 unique edges including: 35,197 friends and followers edges, 14,428 tweet edges, 358 mention edges, and 75 reply edges. We studied their network structure in an attempt to understand how they operate and compare how different they are from the Crimean water crisis bots. As depicted in Figure 12.7, we found out the bots network here is not as simple as the one in the Crimean water crisis. Bots here do not follow the principles of "Follow Me and I Follow You" (FMIFY) and "I Follow You, Follow Me" (IFYFM) but the identification of these bots has been challenging because the bots here are also coordinating. This behavior is more pronounced in the communication network (retweet + mention + reply). Meaning, if we look at the friends–followers network we don't see much coordination (unlike the Crimean water crisis simple bots). While looking at the communication network it does reflect coordination, and that too is observed by applying a very sophisticated network structure analysis algorithm, i.e., our FSA approach [5].

In the Trident Juncture Exercise, we did an empirical study of the anti-OTAN and anti-TRJE 2015 cyber propaganda campaigns organized on Twitter and the blogosphere. Domain experts identified six groups that propagated their messages on social media inviting people to act against NATO and TRJE 2015. We identified their blog sites as well as their Twitter handles using Google search and cyber forensics techniques. Then we collected their Twitter network for the period of August 3, 2014 4:51:47 PM UTC to September 12, 2015 3:22:24 AM UTC. This

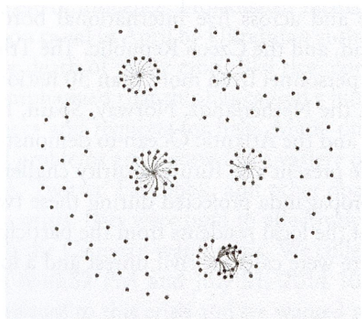


Figure 12.7 Coordination among Dragoon Ride (2015) bots discovered on examining the communication network. Social network does not exhibit any coordination.

resulted in 10,805 friends–followers, 68 replies, 654 tweets, 1,365 mentions, 9,129 total unique nodes, and 10,824 total unique edges.

Then we used cyber forensics tools (Maltego: an open source intelligence and forensics application. It saves a lot of time in mining and gathering of information as well as the representation of this information in an easy to understand format [see Figure 12.8]) and cyber forensics techniques, such as tracking Google Analytics ID, which is an online analytics tool that allows a website owner to gather some statistics about their website visitors such as their browser, operating system, and country. As depicted in Figure 12.8, multiple sites can be managed under a single Google analytics account. Using the techniques in [8] and [9], we were able to uncover the hidden relations between different blog sites and also to study the cross media affiliation of different groups.

After identifying these blogs, we crawled them and fed their data to our in-house developed Blogtrackers tool, where further analysis on the blogs level as well as blogger level can be conducted. Using Blogtrackers we were able to see the spike of activity of blogs (number of blog posts) just before the Trident Juncture Exercise (see Figure 12.9). We were also able to identify the most influential posts [17] that happened to have a lot of propaganda and a clear call for civil unrest against NATO forces (see Figure 12.10).

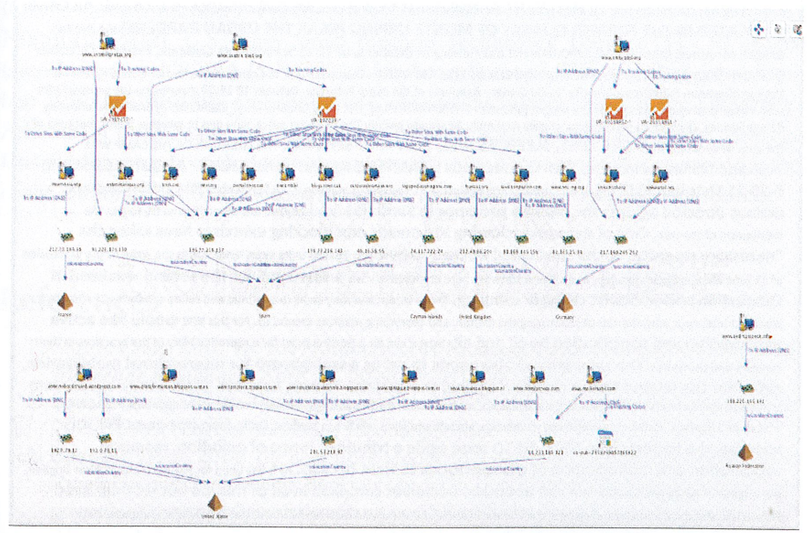


Figure 12.8 Finding related websites, IP addresses, and locations of websites using the Maltego tool and the unique identifier (Google Analytics ID).

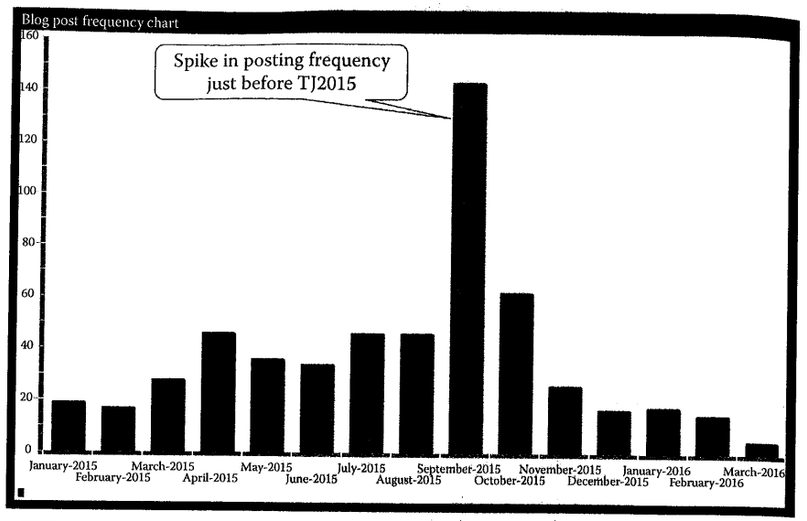


Figure 12.9 This shows the spike in blogs posting activity just before the Trident Juncture Exercise.

camp program: TODAY October 11, 2015 **EVENT Antimilitarist** TO 18 IN CAGLIARI. CONCENTRATION PIAZZA D'ARMI. THE **CAMP IS LOCATED IN THE FORMER QUARRY OF MONTE URPINU (NEAR THE URBAN GARDENS)** VIA RAFFA GARZIA. For visitors REMAINS THE APPOINTMENT this morning on October 9, UP TO 11 IN PIAZZA DEL CARMINE. **Friday, October 9: From 9 to 11 reception in the square of the carmine** - the opening of the camp in the former quarry at Monte Urpinu- afternoon initiatives in the city '21.00 dinner - Assembly of the camp Saturday, October 10 18:00 meeting on the prospects for anti-militarist struggle and against the trident juncture - PRESENTATION of THE NEW CALENDAR of EXERCISES IN SARDINIA following dinner Sunday, October 11 Morning conclusive Assembly Afternoon parade The program potra 'vary due to weather issues, because of the cops or contingency. **PORT TENT, SLEEPING BAG, FLAT AND SERVERS.** THE LOCATION OF THE CAMP WILL 'PUBLISHED TOMORROW MORNING, THEN Meet RECEPTION II **CAMPING Antimilitarist FIGHT - AROUND CAGLIARI 9-10-11 October 2015** Out of the mobilization against the Capo Frasca polygon of 13 September 2014, **initiatives and actions directed against the military presence in Sardinia** have multiplied and diversified to try to jam the mechanism of the war. **Cuts of networks, slowing the means and blocking exercises have taken the "necessary serenity" to the conduct of military activities.** Thanks to its experience and in the wake of the procession of 11 June 2015 in Decimomannu, as No Bases Network here or elsewhere we decided to **call for the second weekend of October an anti-militarist struggle camping.** These three days they want to continue and refine the forms of struggle practiced until now, with the aim of sabotaging the military and everything revolves around us. For this **we would like active participation and contribution by all and all,** then it can be a starting point for a reproducibility of the practices in their contexts and territories. **The campground also wants to act as a springboard for international mobilization, called for the second half of October, against the exercise Trident Juncture 2015.** With this exercise, NATO intends to test its intervention force in the short term, to prepare for the increasingly Possible conflict on Middle East fronts, North African and Russian. 36000 men, hundreds of vehicles, aircraft and ships will fire in Sardinia, Sicily, Spain and Portugal. **For this exercise, the largest since 2002, NATO once again a tribute in terms of pollution, resource exploitation and militarization of the territories to train for war.** As it has been for the exercises of Aries brigade, the brigade of Aosta and STAREX, **we can not make ourselves complicit in all of this, do not let them rest assured.** Proposal mobilization against the Trident DOWNLOAD INFORMATION MEMORANDUM ON TRIDENT Juncture 2015 46

Figure 12.10 This shows the most influential post, which has a clear call for civil unrest against the NATO forces.

12.5 Conclusion and Future Work

In conclusion, with the rapid advancement of technology, people are more connected than ever before. Internet and social media boosted the speed of information diffusion tremendously across the globe. Thus, disseminating propaganda or misinformation about events, i.e., conducting deviant acts, becomes convenient, effective, and fast. Deviant groups can nowadays coordinate cyber campaigns in order to achieve strategic goals, influence mass thinking, and steer behaviors or perspectives about an event in a very sophisticated way (hard to discover) and yet easy to be done. In this chapter, we provided two important and detailed case studies, namely Daesh (ISIS: Islamic State in Iraq and Syria/ISIS/ISIL) and Novorossiya. We analyzed the situational awareness of the real world information environment in and around those events by employing computational social network analysis and cyber forensics informed methodologies to study information competitors who seek to take the initiative and the strategic message away from the main event in order to further their own agenda. We showed the methodology we followed and the results we obtained from each case study. The research gave many interesting findings that are mentioned above and was of great benefit to NATO and U.S. forces participating in both exercises on the ground.

Acknowledgments

This research is supported in part by grants from the U.S. Office of Naval Research (ONR) (award numbers: N000141010091, N000141410489, N000141612016, and N000141612412), U.S. Navy Small Business Technology Transfer (STTR) program (award number: N00014-15-P-1187), U.S. Army Research Office (ARO) (award number: W911NF1610189), and the Jerry L. Maulden/Entergy Fund at the University of Arkansas at Little Rock. The researchers gratefully acknowledge the support. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organization.

References

1. Smith, Craig. 2016. By The Numbers: 170+ Amazing Twitter Statistics. *DMR (Digital Marketing Ramblings)*. April 30. http://bit.ly/1bSfjNi.
2. Smith, Craig. 2016. By The Numbers: 200 Surprising Facebook Statistics (April 2016). *DMR (Digital Marketing Ramblings)*. June 1. http://bit.ly/1qVayhl.
3. Sindelar, Daisy. 2014. The Kremlin's Troll Army: Moscow Is Financing Legions of pro-Russia Internet Commenters. But How Much Do They Matter? *The Atlantic*, August. http://www.theatlantic.com/international/archive/2014/08/the-kremlins-troll-army/375932/.
4. Al-khateeb, Samer, and Nitin Agarwal. 2015. Examining Botnet Behaviors for Propaganda Dissemination: A Case Study of ISIL's Beheading Videos-Based Propaganda. In *Data Mining Workshop (ICDMW), 2015 IEEE International Conference on*, 51–57. IEEE.

5. Sen, Fatih, Rolf Wigand, Nitin Agarwal, Serpil Yuce, and Rafal Kasprzyk. 2016. Focal Structures Analysis: Identifying Influential Sets of Individuals in a Social Network. *Social Networks Analysis and Mining* 6: 1–22. doi:10.1007/s13278-016-0319-z.
6. Rodríguez-Gómez, Rafael A., Gabriel Maciá-Fernández, and Pedro García-Teodoro. 2013. Survey and Taxonomy of Botnet Research through Life-Cycle. *ACM Computing Surveys (CSUR)* 4(4): 45.
7. Abokhodair, Norah, Daisy Yoo, and David W. McDonald. 2015. Dissecting a Social Botnet: Growth, Content and Influence in Twitter. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 839–51. ACM. http://dl.acm.org/citation.cfm?id=2675208.
8. Alexander, Lawrence. 2015. Open-Source Information Reveals Pro-Kremlin Web Campaign. News Website. *Global Voices*. July 13. https://globalvoices.org/2015/07/13/open-source-information-reveals-pro-kremlin-web-campaign/.
9. Bazzell, Michael. 2014. Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information. 4th ed. CCI Publishing. https://inteltechniques.com/book1.html.
10. Al-khateeb, Samer, and Nitin Agarwal. 2014. Developing a Conceptual Framework for Modeling Deviant Cyber Flash Mob: A Socio-Computational Approach Leveraging Hypergraph Constructs. *The Journal of Digital Forensics, Security and Law: JDFSL* 9 (2): 113.
11. Al-khateeb, Samer, and Nitin Agarwal. 2015. Analyzing Deviant Cyber Flash Mobs of ISIL on Twitter. In *Social Computing, Behavioral-Cultural Modeling, and Prediction*, 251–57. Springer.
12. Carter, Joseph A., Shiraz Maher, and Peter R. Neumann. 2014. #Greenbirds: Measuring Importance and Influence in Syrian Foreign Fighter Networks. The International Center for the Study of Radicalization and Political Violence (ICSR). http://icsr.info/wp-content/uploads/2014/04/ICSR-Report-Greenbirds-Measuring-Importance-and-Infleunce-in-Syrian-Foreign-Fighter-Networks.pdf.
13. DeHaan, Mike. 2012. Introducing Math Symbols for Union and Intersection. *Decoded Science*. July 26. http://www.decodedscience.org/introducing-math-symbols-union-intersection/16364.
14. Al-khateeb, Samer, and Nitin Agarwal. 2016. Understanding Strategic Information Manoeuvres in Network Media to Advance Cyber Operations: A Case Study Analysing pro-Russian Separatists' Cyber Information Operations in Crimean Water Crisis. *Journal on Baltic Security* 2(1): 6–17.
15. Ghosh, Saptarshi, Bimal Viswanath, Farshad Kooti, Naveen Kumar Sharma, Gautam Korlam, Fabricio Benevenuto, Niloy Ganguly, and Krishna Phani Gummadi. 2012. Understanding and Combating Link Farming in the Twitter Social Network. In *Proceedings of the 21st International Conference on World Wide Web*, 61–70. ACM. http://dl.acm.org/citation.cfm?id=2187846.
16. Labatut, Vincent, Nicolas Dugue, and Anthony Perez. 2014. Identifying the Community Roles of Social Capitalists in the Twitter Network. In *The 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, China. doi:10.1109/ASONAM.2014.6921612.
17. Agarwal, Nitin, Huan Liu, Lei Tang, and Philip S. Yu. 2012. Modeling Blogger Influence in a Community. *Social Network Analysis and Mining* 2(2): 139–62.

III TOOLS AND DATASETS FOR CYBERSECURITY