Countering Daesh Cognitive and Cyber Warfare with OSINT and Basic Data Mining Tools

Gianluigi ME1, Maria Felicita MUCCI2
1 Department of Economics and Finance, Luiss Guido Carli University, Rome, Italy
gme@luiss.it
2 S & A | Sistemi & Automazione S.p.A., Rome, Italy
mfmucci@sealink.it

Abstract

Digital civilization has changed war circumstances. Emerging dangers have asymmetry, variety, and continual change; quick transmission through the network; near-immediacy; possibility for unrestricted access; and swift power to affect people’s behavior. Cognitive Warfare, an international relations issue, uses information, cyber, and psychological warfare tactics. Daesh sends threatening messages to Western countries and spreads internet propaganda to recruit new members and induce terror. The study attempts to propose a novel knowledge-based approach for detecting terrorists by examining data obtained from Twitter and leading Daesh publications, through Data Mining techniques.

Index terms: clustering analysis, cognitive warfare, counterterrorism, Daesh, national security, virtual jihad

1. Introduction

International power shifts since the Cold War have pushed us closer to a multipolar world. Top-down and bottom-up processes spawn new actors and security trends. We no longer face threats that can be spatially contained in an attack by a big power against another. International terrorism and cybercrime have joined military challenges in national security discussions. Though interrelated, these new structural trends are fragmented and multidimensional [1]. Today’s threats are asymmetrical, diverse, and ever-changing, spread over the network, are immediate, may be publicly available, and can easily impact behaviour. New digital technologies and social media have allowed actors to reach a wider audience with tailored and targeted content. Hostile propaganda players are aware of the cyber environment’s prospects, and they are working constantly to exploit information and undermine its essential truth principles. In addition to providing more access to information, the Internet offers greater opportunities for deception. Indeed, Cognitive Warfare is becoming a crucial aspect in international politics and a growing source of concern.

The current historical era, highlighted initially by the Covid-10 Pandemic crisis and finally by the outbreak of the Russian-Ukrainian conflict, has redirected media and major international powers’ focus away from the terrorist threat. Despite this apparent stalemate, Daesh keeps on operating and establishing itself using many technologies to carry out its acts and plans. It has initiated disruptive cyber and cognitive warfare efforts, including Cyber-Training, Cyber-Planning and Cyber-Execution, Funding and Fundraising, conducting Cyber Attacks, recruiting new members, and disseminating online media propaganda. Virtual Jihad is becoming an increasingly alarming issue that must be addressed.
Jihadism is a relatively new phenomenon in terms of both its objectives and means. Due to technical advancements, the communications revolution, and improvements in information storage and retrieval, previously unthinkable techniques for bringing a community to the forefront of people’s thoughts have developed. The current generation of Jihadists is the first to grow up with pervasive access to digital communication tools. This makes them the first generation capable of joining and directing terrorist organizations. So, it is unsurprising that these networks are essential to the radicalization and the recruiting techniques they employ to attract vulnerable individuals. Jihadists have demonstrated their proficiency at utilizing globalization’s resources to achieve their own objectives through this strategy. Social media platforms and the Internet have become a highly effective tool for spreading propaganda, instigating violence, and radicalizing a far larger audience than in the past, and - if managed properly – they may become powerful psychological weapons with disruptive effects. Indeed, we now refer to the Weaponization of Media Narratives: the battle of narratives has surpassed the conventional military and physical Jihad in importance.

Thus, it is essential to design and deploy Cyber Defense methods to prevent, identify, and dissuade jihadist Internet activity. In this environment, law enforcement, intelligence, and other agencies are always inventing new techniques to prevent, identify, and limit terrorist activities on the Internet. Due to their effectiveness in promptly identifying possible terrorist threats, traditional research techniques are gaining favour again. Moreover, in the last several decades, the collection and analysis of data from a broad variety of sources, in addition to text analysis, has been able to give intelligence analysts with useful insights by revealing previously hidden yet logically sound patterns and connections. Furthermore, Open Source Intelligence (OSINT) and Social Media Intelligence (SOCMINT) are employed to collect pertinent data useful for proving added information about threats. Then, with a fundamental comprehension of Machine Learning (ML) techniques, one may construct a model from a collection of inputs and use it to generate predictions or judgments.

This study aims to develop an innovative knowledge-based technique that may be utilized to detect jihadist affiliation by analysing data collected from social network platforms (particularly Twitter) and the leading Daesh publications. The suggested method may determine the conventional behaviour (or “profile”) of militants and sympathizers because it employs a data mining approach, notably clustering analysis, to analyse digital content associated with Daesh.

2. Virtual Jihad

If we now refer to the modus operandi of Daesh, particular attention should be paid to the way online tools are used, as they have become one of the fundamental elements of terrorists’ activities [2]. According to Weimann, there are two distinct categories to which Jihadists’ use of the Internet can be classified: communicative and instrumental [3]. The first categorization includes spreading propaganda, running psychological warfare campaigns, and recruiting new members, while the second one includes cyber-training, cyber-planning and coordination, and digital fundraising.

Daesh has become increasingly reliant on the Internet in the last decades as a substitute for traditional training grounds. The face-to-face nature of certain activities has given way to their virtual counterparts in modern times. “Virtual training camps” [4] provide a platform for prospective recruits to learn about and support terrorist groups, while also encouraging involvement in direct acts by providing access to encryption tools and anonymizing tactics. Weimann claims that the readily available multimedia format in several languages has become a “terrorist university,” a place where jihadists may learn new strategies and abilities that make their assault methodology more efficient. Because of the Internet’s inherent interactivity, individuals from all over the world can feel connected to one another and form networks to share strategies and tactics. Indeed, there is a wealth of information online on how to join terrorist groups, how to plan and execute terrorist acts, and how to construct explosive devices and other weapons (Figure 1).
In recent decades, nearly every terrorist attack has involved the use of the Internet and the opportunities it provides [5]. Typically, the planning of a terrorist act requires communication between multiple parties over greater distances and utilizing more sophisticated methods. Daesh’s use of anonymizing communication tools to plot attacks has progressed to a new level of sophistication. Tools for encrypting data and software designed to mask a user’s identity make it difficult to determine the sender, the addressee, and - most generally - the contents themselves [6].

Not only does a Jihadist organization require a variety of resources to finance its operations, but also to maintain its existence and grow as an organization. Funds are required for a variety of reasons, among all to maintain militants and their relatives; to finance travel expenses; to recruit and train new members; to acquire weapons and safe houses; to carry out operations; and to promote the ideology of the group through social activities or propaganda. So, it is crucial for the success and resilience of the groups to have access to financial resources, particularly at the beginning, when these are required to assist recruit and maintain support, as well as to create major material capabilities [7].

Moreover, one sort of disruptive cyber intelligence activity that violent radical political parties engage in is the dissemination of online propaganda information designed to attract as many people as possible, and to recruit new members. Communication geared at terrorist recruiting is typically designed with the intention of appealing to weak and marginalized subgroups within society [8]. Consequently, the efficacy of this propaganda in terms of recruitment and radicalization depends on an individual’s sentiments of injustice, alienation, or guilt. Potential terrorist recruits and existing members of the terrorist organization can develop a type of virtual community through interactive engagement, which can also promote a sense of belonging and bolster a sense of community.

In the battle for the Umma’s affections, narrative warfare has surpassed the employment of conventional firearms and military might. The offensive information warfare focused on propaganda is a key component of Daesh’s struggle. Therefore, media and Internet resources may be a powerful psychological weapon if handled appropriately. This narrative-driven, intensified kind of terrorism has emerged as Daesh’s primary asymmetric weapon.

By using a wide range of publicly accessible social networking channels, terrorist and insurgent organizations today have established an even more direct and personally intimate method of message. Some examples of these platforms include Twitter and Telegram amongst others. An example of a potential Twitter account affiliated with Daesh, but already suspended, can be seen in Figure 2.

**Fig. 1.** A Tweet of a suspected Jihadist account, who refers to the “amazing” military infographics of Al-Naba’ Magazine
According to Wilkinson, “When one says terrorism in a democratic society, one also says media. For terrorism by its very nature is a psychological weapon which depends upon communicating a threat to the wider society. This is why terrorism, and the media enjoy a symbiotic relationship” [9]. Distinctly poignant is the analysis conducted by Adam Chuipka for the University of Ottawa: “Media coverage and terrorism are soul mates - virtually inseparable. They feed off each other. They together create a dance of death - the one for political or ideological motives, the other for commercial success” [10].

Daesh is at the frontline of a new revolution in jihadist communication because of the remarkable efficacy with which it uses these platforms to address a worldwide audience. The group will continue to use mass media to get publicity and support because of the glamour of cutting-edge tech and the incredible speed and reach of person-to-person power enabled by modern technology. Given the capabilities and products that are certain to become more advanced in terms of quality, content, speed, affection, and transmission capacity, as well as more numerous and pervasive than ever before, it is possible that we are only now beginning to comprehend the implications of this phenomenon.

3. Weaponization of Media Narratives

Jihadist groups recognize the importance of online media platforms in their strategic planning, as seen by their own attempts to embrace the media requirement. The war of narratives has taken a more prominent role than the traditional use of guns and fire weapons in the conflict that is taking place for control of the hearts and minds of the umma. Members of Daesh ideological group typically identify themselves primarily in contrast to members of other groups, making a significant separation between them and whom the others are (Us vs them). This dualism is so important to the process of forming the Islamist identity. A group’s internal cohesion and belief in its own mission may benefit from this, and the group’s success might inspire others to show support for or even join the group. On the other hand, the same propaganda may instil panic in the minds of people who have been singled out for terrorist attacks.

Da’wah (Arabic: دعوة) is an essential part of the Daesh’s ideology. Its purpose is to recruit, indoctrinate, and motivate terrorist attackers, supporters, and sympathizers. Active Da’wah is carried out in online social networks as well as on the websites of the various Islamist and Jihadi organizations, with the aim of accomplishing four important goals: informing, frightening, uniting, and supporting [11]. The jihadists’ timely and impressively well-informed posts in social networks
and blogs, as well as comments on Islamism websites, reveal their knowledge with daily political news, and political decision-making processes made my Western countries. This approach makes it clear that a well-articulated media strategy is in no way haphazard, but rather makes use of skills taken from the realm of communication that is globally pervasive and widespread. Terrorists may benefit in a number of ways when they use the Internet to spread their messages: they can remain hidden from public scrutiny, they can quickly and easily reach a wide audience, and they can encourage dialogue amongst their followers.

Twitter is one of the social networks that might be leveraged to accomplish this goal. Supporters of terrorist organizations may send and receive messages, photographs, videos, and website connections to a broad audience via Twitter. The platform acts as a venue for both passive and active supporters. It is a possible danger due to its ability to distribute instant messages to a huge number of users simultaneously and because it allows users to follow specified subjects and groups, as well as other users’ tweets about those topics. Indeed, Daesh uses Twitter as an umbrella platform, which combines the numerous information sources into a unified index (mostly via the use of hashtags) that can be viewed and searched with relative ease. The weekly magazine Al-Naba’ (which started its activity on 17 October 2015 to now) is an essential Daesh propaganda publication that we may consider in our discourse. While all Dabiq and Rumiyah magazines are published in English, Al-Naba’ issues are only published in Arabic. The Magazine is published on a regular basis by the Diwan al-Ilam al-Markazi, which is the central media organization for Daesh, and it is responsible for coordinating media efforts and providing guidance to its media supporters. Al-Naba’ features a variety of articles and material types, such as news, commentary, visualizations, religious writings (including fatwas), and advertisements for various forms of media output. So, it reflects Daesh mindset and mirror events occurring on the ground, and social issues particularly relevant to the supporters.

4. Countering Terrorist through OSINT, SOCMINT, and Data Mining Tools

It is not enough for governments to only explore new ways to acquire and combine Intelligence; they must also analyse it to provide actionable information in the fight against Daesh. Open-Source Intelligence (OSINT) is an important discipline for the processing of publicly accessible sources. It has become an increasingly valuable source of intelligence for governments, companies, and criminals alike, as more and more detailed personal information is processed digitally and accessible online. Further, a variety of searches, including picture searches, web text investigations, social media content searches, and map searches, are all accessible [12]. Social Media Intelligence (SOCMINT), a subset of OSINT, is the study of how social networking services can be monitored and mined for useful information and community detection and analysis.

The collection and analysis of information from wide variety of sources can provide intelligence analysts with relevant insights as they can reveal previously hidden yet logically sounds patterns and linkages. OSINT and SOCMINT may be interchangeably used to collect relevant data valuable for providing further information about specific threats. A basic understanding of Machine Learning (ML) algorithms can make a mathematical model from a set of inputs and utilize it to make predictions or judgements.

4.1. Aim of our study

Our research aimed to determine whether Machine Learning (ML) and Natural Language Processing (NLP) can be used to analyse Jihadist stories in order to find any similarities between different sources of propaganda. One of the specific goals was to evaluate whether or not there are tweets with a direct connection to Al-Naba’ magazine. The volume of propaganda released by Daesh is so large that it is nearly impossible for humans to analyse it. As a result, establishing methods and procedures that can be utilized to analyse massive amounts of data is a crucial challenge. The
development of counter-narratives and counter-messaging strategies requires an understanding of the various taxonomies used in propaganda, the ways in which the narrative varies across media outlets, and the way in which it develops over users. *Social Network Analysis* and *Data Mining* – specifically *clustering algorithms* – where applied in order to achieve our main objectives. Finally, Tweets collected from potential Daesh supporters and sympathizers were used as a target of our data investigation.

### 4.2. Research Objectives

Examining Daesh propaganda narrative and the most common taxonomies in *Al-Naba’* magazine.

Empirical search for a tiny similarity algorithm.

Identifying Daesh affiliations to *Al-Naba’* through social media analysis of Twitter using Data Mining Techniques, in particular clustering.

### 4.3. Justification

The goal of this study was to develop a model that would aid in the detection of *Al-Naba’* Daesh affiliation via the use of clustering algorithms. As a result, the proposed model would offer a framework for identifying jihadist hidden propaganda with little human intervention. This attempts to improve security organs by providing a more precise and quick way of detection than previous techniques.

### 4.4. Research Phases

Our empirical research started from the collection of the most relevant issues of *Al-Naba’* newspaper, published on a regular basis by the *Diwan al-Ilam al-Markazi*. *Al-Naba’* features a variety of articles and material types, such as news, commentary, visualizations, religious writings (including fatwas), and advertisements for various forms of media output. So, it totally reflects Daesh mindset and mirror events occurring on the ground, and social issues particularly relevant to the sympathizers. The articles were selected for their relevance as being particularly under the attention of the Western world’s monitoring and inspection instruments. The reason we analyzed *Al-Naba’* magazine rather than another newspaper is because it has only been published in Arabic, whereas other journals have been released in other languages and have been the subject of numerous prior studies. However, no study has been conducted on *Al-Naba’*. Here is the first difficult encountered: a literal translation was adopted in order to diverge as little as possible from the expressive tactics and features used by Daesh adherents. A previous careful study of *Dabiq* and *Rumiyah* magazines (published in English) offered key insights for the overall translation. Several articles of each issue were then selected randomly, and their translation was refined.

The subsequent phase entailed the identification of taxonomies to be extracted for our investigations that appeared frequently. The statements were incorporated into Expert.AI’s Cogito Intelligence API (an online semantic-based system dedicated to crime and intelligence that also includes specialized classification) which offers full semantic processing patterns like Categorization, Text Mining and Fact Mining. *Al-Naba’* extracts were selected and sequentially placed in the section to be analysed in order to determine the most pertinent metadata. Consequently, we initiated the creation of our dataset in a shared Excel file. Its line-column chart is a combination of a line graph (containing *Al-Naba’* texts) and a set of columns that correspond to the extracted Cogito system parameters, with the fewest possible empty cells. Using a JSON score, we formed a dataset in which each text was defined by a vector (a list) of scores from all categories present in all documents. The dataset was balanced by converting the scores (through such that instead of having scores ranging from 0 to infinity, they now have values ranging from 0 to 1 (Figure 3). In order to do so, we used sklearn.preprocessing.MinMaxScaler.
We fed this dataset to the clustering algorithm by choosing a number of five clusters (Figure 4).

Therefore, a dataset was generated with the column at the end. We tried to discover the correlations between the clusters given to each text and the categories activated. At the end we noticed two clusters concentrated on the category explosion/explosives and terrorism/act of terror. Once we had trained the clustering model based on n text categories, we investigated whether feeding Tweets were meaningfully associated with the categories. Several OSINT methods were used to extract tweets from users that may have a particular relation with Daesh and the Jihadist propaganda. First, we used Twitter Advanced Search with the following query:

داعش
الدولة الإسلامية
الدولة الإسلامية
IPS
ISIS
OR
Twitter
until:2022-11-09
since:2022-01-01

Among the search results, one account in particular (@a_o_be_dh90) caught our curiosity, but it had been suspended within a few days (Figure 1).

Our initial investigation included not just the content of published posts, but also followers and followings. In terms of saving time and do more accurate research, we decided to employ another OSINT tool, namely TweetTopic. Such instrument offers a function that we have found useful in our investigation. Once the Twitter identity is provided, the tool gathers the most recent 3,000 Tweets and generates a word cloud. This detects the most frequently used terms inside the target’s postings. When someone clicks on a phrase among the search results, only Tweets containing the specified terms are displayed. When you click on any of the text circles, the corresponding postings are immediately shown. It was incredibly useful when we had too many posts to read in a short amount of time. TweetTopic enables us to swiftly determine the content of our target’s tweets and to promptly investigate any relevant subjects. In addition, the word cloud displays the most frequent retweets or ID accounts associated with the evaluated tweet. As a result, a set of 100 Twitter accounts (ten of them are shown in Error! Reference source not found.) Table 1 and potentially related to Daesh – 93 of them created from February 2022 to 8 November 2022 – was formed. Some of their tweets contain mentions to articles taken from Al-Naba’ magazine (Figure 3).

Table 1. Suspected Daesh Accounts on Twitter

<table>
<thead>
<tr>
<th>ACCOUNT</th>
<th>LINK</th>
<th>REGISTRATION TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Account-name_1</td>
<td>twitter-url-account-name_1</td>
<td>September 2022</td>
</tr>
<tr>
<td>Account-name_2</td>
<td>twitter-url-account-name_2</td>
<td>September 2022</td>
</tr>
<tr>
<td>Account-name_3</td>
<td>twitter-url-account-name_3</td>
<td>September 2021</td>
</tr>
</tbody>
</table>
TABLE 1: ACCOUNT, LINK, REGISTRATION TIME

<table>
<thead>
<tr>
<th>ACCOUNT</th>
<th>LINK</th>
<th>REGISTRATION TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Account-name_4</td>
<td>twitter-url-account-name_4</td>
<td>February 2018</td>
</tr>
<tr>
<td>Account-name_5</td>
<td>twitter-url-account-name_5</td>
<td>September 2012</td>
</tr>
<tr>
<td>Account-name_6</td>
<td>twitter-url-account-name_6</td>
<td>October 2022</td>
</tr>
<tr>
<td>Account-name_7</td>
<td>twitter-url-account-name_7</td>
<td>September 2022</td>
</tr>
<tr>
<td>Account-name_8</td>
<td>twitter-url-account-name_8</td>
<td>November 2022</td>
</tr>
<tr>
<td>Account-name_9</td>
<td>twitter-url-account-name_9</td>
<td>July 2022</td>
</tr>
<tr>
<td>Account-name_10</td>
<td>twitter-url-account-name_10</td>
<td>October 2022</td>
</tr>
</tbody>
</table>

Fig. 5. Tweets, on Al-Naba’ Magazine, of suspected Daesh-related accounts

Assuming that the tweets found could be classified into the identified clusters, we attempted to code them in the same manner as previously applied with the existing documents to see whether they were classified in the cluster in which we had expected them to be categorized.

4.5. Research Scope and Limitations

Obviously, our research has limits, there are still many unknown concerns. Neither the work nor the analysis is adequate to perform a complete inquiry. Notwithstanding these drawbacks, we sought to illustrate the ability of complex networks to reveal hidden patterns within the global context of terrorism to stimulate the use of fresh analytical frameworks to the study of Daesh media narratives. By finding hidden patterns across several social and digital channels, similarities in propaganda can be revealed. According to our assessment, there is an urgent need to develop and evaluate innovative methodological techniques that, if effective, can be applied to other contexts with more precise data, allowing for additional useful and decisive conclusions.

5. Conclusions

Daesh keeps on using several technologies to carry out its deeds and plans. Virtual Jihad is becoming more and more an issue of concern that must be addressed and it depends largely on offensive propaganda-based information warfare. This narrative-driven, heightened form of terrorism has become a major asymmetric weapon. Online terrorism, specifically, has long-lasting effects on the psyche (“Cognitive Warfare”) of impacted communities and virtual users and is capable of causing significant harm.
The fight against the jihadist threat has had some success over the last two decades, but in the next years, it will need to take more into consideration the hazards that progress on the World Wide Web. As a result, Law Enforcement, Intelligence Agencies, and other organizations must constantly develop innovative unique techniques to prevent, identify, and limit terrorist activities over the Internet. Intelligence gathered and evaluated will aid in assessing threats, formulating responses, and guiding policies. In this process, the influence of new technologies and social shifts on the operation of Law Enforcement must be considered. Additionally, cultural, and digital awareness are essential for fostering State’s collaboration and partnership in the fight against cyberterrorism.

Therefore, with our study, we aimed to demonstrate that relatively simple data mining tools with social network analysis features, when combined with conventional data mining techniques and practical semantic analysis of online propaganda, can serve as a useful starting point for identifying terrorist affiliations. Our approach began with the following question: is it possible to employ Machine Learning and Natural Language Processing algorithms to assess Daesh narratives in order to identify possible similarities among different propaganda sources?

The volume of propaganda disseminated by Daesh is so big that it is nearly difficult for human capabilities to examine it. As a result, developing methods that can be applied to analyse massive amounts of data was a key task. Our elaboration of counter-narratives and counter-messaging approaches required a basic knowledge of Arabic language, an understanding of the several taxonomies used in propaganda, a consideration of ways in which the narrative varies among media channels and the way it develops over users. Social Media Analysis and Data Mining tools – specifically clustering algorithms – where applied to achieve our main objectives. Finally, Tweets collected from potential Daesh supporters and sympathizers were used as a target of our data investigation.

Given the employed methodology, there is no reason to doubt that – despite the potential drawbacks – there are additional potential avenues for future research.

References


hope that your project is developing well. Can we get some update on published results/attachment/59d64a2479197b80779a47e9/AS:474077556154368@1490040305737/download/ZbornikN20015.pdf#page=208.

[7]. J. Adams, “The financing of terror: Behind the PLO, Ira, red brigades and M-19 stand the paymasters: how the groups that are terrorizing the world get the money to do it,” New York: Simon and Schuster, Jan. 1986.


