

Measuring the Impact of ISIS Social Media Strategy

Majid Alfifi, Parisa Kaghazgaran, and James Caverlee
Department of Computer Science & Engineering
Texas A&M University
College Station, TX
{alfifi, kaghazgaran, caverlee}@tamu.edu

Fred Morstatter
Information Sciences Institute
University of Southern California
Marina Del Rey, CA
fredmors@isi.edu

ABSTRACT

Terrorist groups like the Islamic State of Iraq and Syria (ISIS) have exploited social media such as Twitter to spread their propaganda and to recruit new members. In this work we study the extent to which ISIS is able to spread their message beyond their immediate supporters. Are they operating in their own sphere with limited interaction with the overall community? Or are they well rooted among normal users? We find that three-quarters of the interactions ISIS received on Twitter in 2015 actually came from eventually suspended accounts raising questions about the potential number of ISIS-related accounts and how organic ISIS audience is. Towards tackling these questions, we have created a unique dataset of 17 million ISIS-related tweets posted in 2015. This dataset is available for research purposes upon request.

1 INTRODUCTION

The past few years have seen social media as an effective tool for facilitating uprisings and enticing dissent in the Middle East [4][11][14]. The embrace of social media in the region has made it a battle ground for ISIS (and similar groups) versus existing regimes, all spreading propaganda, recruiting sympathizers, and undermining rivals [4]. Social media has given terrorists the ability to directly come into contact with their target audience and either spread terror or recruit. In fact, ISIS has been repeatedly described as the most adept terrorist group at using Internet and social media propaganda to recruit new members [13]. Several studies have looked into how ISIS operates on platforms like Twitter [3] [8] [12] but little has been done to quantify their true reach and impact. While there are many factors that could have contributed to their success – such as gaining territories on the ground and appealing to the communities where they operate [12] – quantifying the impact of their social strategy is still crucial for a better understanding of their operations. In addition, although Twitter suspends accounts that violate the terms of service, ISIS accounts seem to succeed in staying longer on the service posting more tweets than other malicious eventually suspended accounts (Figure 1) – making studying their impact even more important.

As we discuss in more detail in Section 3, several ISIS-related datasets have been made available to the research community. Our

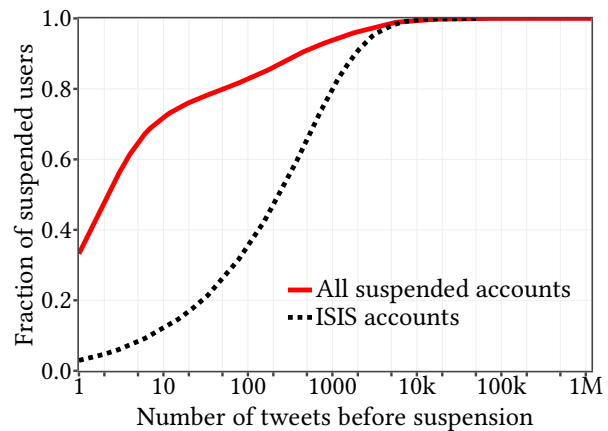


Figure 1: ISIS accounts spread more content before getting suspended by Twitter compared with other eventually suspended accounts.

work complements those efforts by contributing a more comprehensive dataset spanning the year 2015. Concretely, we share with the research community a dataset of 24k ISIS users and their 17 million tweets that span all of 2015. We additionally make an initial look into the impact of their behavior on the Twitter community as a whole in Section 4.

2 DATASET

We have access to a large dataset of 9.3 billion tweets representing all tweets generated in the Arabic language in 2015 through private full access to the Twitter Firehose. In addition, we exploit a crowdsourcing initiative by the Anonymous hacking group that invited Arabic speakers to report Twitter accounts that they think were associated with ISIS. This effort originally identified more than 25,000 ISIS sympathizers and supporters through crowdsourced reporting.¹ We only use accounts that have actually been suspended by Twitter, indicating multiple users have reported those accounts and that the Twitter spam control team found those accounts in violation of the terms of service. We then look for their tweets and their interactions with the Twitter community in our larger dataset. A previous work [8] used this same ISIS users dataset but, due to Twitter limitations, they were limited to 10% of the total tweets through the Truthy project at Indiana University [6]. In

¹The original website hosting these accounts has been taken offline but we were able to recover accounts from <http://archive.is/A6f3L>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MIS2, 2018, Marina Del Rey, CA, USA

© 2018 Copyright held by the owner/author(s).

ACM ISBN 123-4567-24-567/08/06.

https://doi.org/10.475/123_4

Dataset	Accounts	Tweets
ISIS-Tweets	23,880	17,434,323
ISIS-Retweets	551,869	10,436,603
ISIS-Mentions	745,721	19,570,380
Legit-Tweets	23,880	17,454,068
Legit-Retweets	1,753,195	12,175,619
Legit-Mentions	2,161,106	17,479,990

Table 1: ISIS-Tweets are tweets posted by a known seed of ISIS-related accounts. Legit-Tweets is a randomly sampled set of users and their tweets. Retweets and mentions of these two sets (ISIS and Legit) by the overall Twitter community are also extracted.

contrast, in this work we were able to recover all of the content these accounts generated in 2015 with our private full access to the Twitter Firehose. Note that this dataset contains only tweets in the Arabic language but since ISIS is mainly active in the Middle East and the majority of its members are Arabic speakers, we believe this dataset provides a clear window into ISIS social media strategy.

To study the impact of ISIS on the overall Twitter community, we focus here on their interactions with the community. We consider *retweeting* as a signal that a user has seen the tweet although they may or may not agree with it. We also consider *reply* or *mention* of a user as an interaction event that often involves discussion, agreement or disagreement. We measure the size of the interaction which is useful in understanding ISIS penetration in the Twitter community. In addition, we make a general observation about the potential sentiment of those who interacted with ISIS based on the keywords used [12].

We select two groups of tweets – one that represents ISIS-related tweets and another randomly sampled group that we use to compare ISIS to other normal users. Concretely, we have the following sets of data (summarized in Table 1):

ISIS-Tweets: Tweets posted by the reported ISIS-related accounts. There are 23,880 accounts that generated 17,424,323 tweets.

ISIS-Retweets: All retweets of ISIS tweets including those from themselves. There are 10,436,603 retweets, posted by 551,869 users.

ISIS-Mentions: All tweets that mention any of the ISIS accounts. There are 19,570,380 such tweets generated by 745,721 accounts.

In order to better understand the behavior and interactions of ISIS accounts, we randomly sampled an equal-size set of legitimate accounts (i.e. accounts that were still alive on Twitter by end of 2016) and use them for comparison. Concretely we have the following:

Legit-Tweets: A random sample of 23,880 legitimate accounts. These accounts posted 17,454,068 tweets.

Legit-Retweets: All retweets of the above legit tweets. We found 12,175,619 retweets generated by 1,753,195 users.

Legit-Mentions: All tweets mentioning the above legit tweets. We found 17,479,990 tweets generated by 2,161,106 users.

3 RELATED WORK

In this section, we first discuss datasets that are relevant to measuring the social media communications of extremists and then we introduce other approaches that have investigated the online communication of extremists.

Berger and Morgan identified 46,000 ISIS supporter accounts on Twitter [4]. They began with 454 accounts known as ISIS supporters as initial seeds. They then collected all accounts following those seeds. This approach continues to two further steps on the followed-by network. They introduced a classification task to determine if a user is an ISIS supporter.

Magdy et al., collected 3 million Arabic tweets referring to ISIS [12]. They classified users based on how they call ISIS. The full name of “Islamic State” is the indicator of support, while abbreviations like “ISIS” or “Daesh” indicate opposition. Zaman [15] introduced an ISIS dataset consisting of about 17,000 tweets from about 100 ISIS supporter accounts. Keywords, Images, and network-based features were used to classify a user as pro-ISIS or not.

In another effort, Bodine-Baron et al., collected 23 million tweets that referenced the Arabic versions of either “Islamic State” or “Daesh” [5]. This corpus consists of both ISIS supporters and opponents whom they were separated by which phrase they use to describe ISIS: ISIS followers refer to ISIS as the “Islamic State” while detractors often use the term “Daesh”. They lexically analyzed the tweets and found that users saying “Daesh” use other terms such as “Terrorist”, “Kharijites”, “militants”, “dogs of fire”, and “dogs of Baghdadi” that prove they are highly critical of ISIS. On the other hand, tweets with “Islamic State” contain glowing terms such as “monotheists Mujahideen”, “Soldiers of the Caliphate”, and “lions of the Islamic State”.

In a recent effort, Badawy et al., investigated how ISIS supporters take advantage of social media to spread their propaganda and recruit militants by studying 1.9 million tweets posted by 25,000 accounts – recognized as pro-ISIS and suspended by Twitter [2].

The Islamic State heavily relies on social media for recruitment. Recently there has been work that focuses on analyzing the data made by this group. For example, Farwell [7] studied the media strategies of ISIS. They described two conflicting strategies followed by ISIS: they try to protect the identity and location of their leaders by minimizing Internet communications while they take advantage of social media for recruitment. Gates and Potter [9] analyzed who creates the content of the organization’s recruitment materials and found that they use a globally distributed network of volunteers who create content to fit the aesthetic of their particular region. Both of these works focus on the recruitment and creation of ISIS messaging. In this work we investigate its effectiveness by measuring the reactions by others on social media.

4 OBSERVATIONS

Are ISIS accounts Spam? We start our analysis by looking into the overall activity of ISIS accounts as compared to other normal users. Do they exhibit a spammy behavior such as posting many tweets or many URLs linking to their own websites? Figure 2 shows that over time ISIS accounts maintain a similar level of activity when compared to other legitimate users although we see that their activity was negatively impacted through the year 2015, due to Twitter closing down their accounts after the crowd started reporting those accounts. In addition, we also see that ISIS accounts manage to post more than other suspended accounts (Figure 1) hinting that they might not have exhibited *banning* signals until later in their lifespan. Moreover, we don’t notice an excessive use of URLs by ISIS

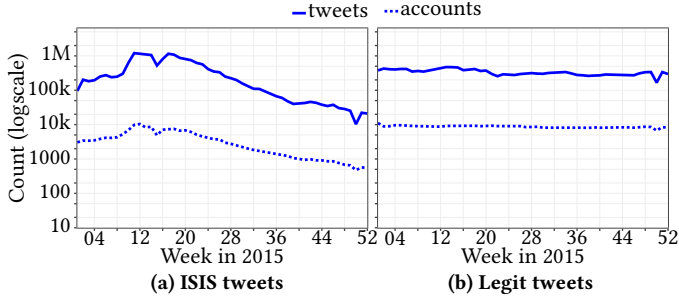


Figure 2: ISIS didn’t post excessively in 2015 (a) compared with a random set of legit users (b). Additionally, ISIS lost accounts over time in 2015 due to Twitter suspending them. The sharp local drop in week 15 and week 50 is due to data collection issues.

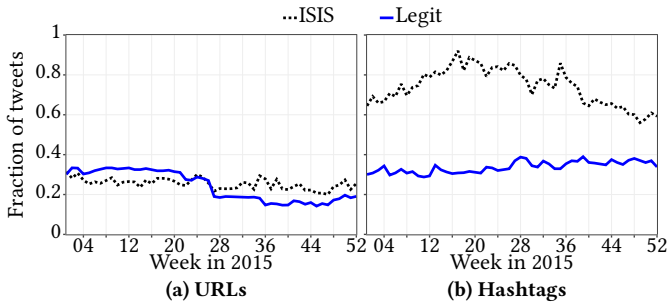


Figure 3: ISIS resembles normal users in their use of URLs (a) clearly different from spam users spreading lots of URLs. However, ISIS uses more hashtags (b), potentially for hashtag hijacking

(Figure 3) clearly distinguishing their accounts from other prevalent spam accounts that post many URLs [10]. Another method by which ISIS may try to reach out to the community is by posting tweets in trending or unrelated hashtags. In fact, we find that ISIS heavily uses hashtags (Figure 3). The reach achieved through this *hashtag hijacking* is difficult to quantify because Twitter users may see their tweets while browsing the hashtag but will not take any action such as retweeting or replying that we are able to collect. For this reason, although important we don’t consider reach obtained through *hashtag hijacking*.

Who retweets/mentions ISIS? Next, we move to studying the first level users who interacted with ISIS accounts. We find that about 76% of the interactions ISIS received actually came from eventually suspended users (Figure 4a). As we reported in [1], 23% of active Arabic Twitter users in 2015 were eventually suspended (contributing 21% to the Arabic Twitter volume that year). This means that, roughly speaking, by mere chance a community of retweeters will be about 20% suspended and 80% not suspended. However, we see that retweets of ISIS content are more than three times likely to be from suspended accounts, strongly suggesting that ISIS retweeters are also involved in malicious activities. The legit group, by comparison, has only 17% of its retweets coming from eventually suspended accounts which is in line with the above 80/20 rule.

Group	Users	Tweets
ISIS Retweeters (Suspended)	170,016	389,358,515
ISIS Retweeters (Not Suspended)	381,853	964,828,227
Total	551,869	1,354,186,742

Table 2: ISIS retweeters

ISIS	ISIS Retweeters (Suspended)	ISIS Retweeters (Not suspended)
Islamic_State	Islamic_State	Saudi
Caliphate_News	Saudi	Decisive_Storm
Caliphate_State	Decisive_Storm	Al-Hilal_FC
Daesh	Caliphate_State	Quran
Saladin_Region	AlHilal_FC	Yemen
Decisive_Storm	Daesh	Syria
Ramadi	Caliphate_News	Hadith
Anbar_Area	AlNasr FC	AlNasr FC
Takrit	Riyadh	Peaceful_Tweeter
Amaq_Agency	Egypt	Egypt

Table 3: Top Hashtags (translated) for ISIS and their retweeters. The top hashtag for suspended retweeters (#Islamic_State) is known to be highly associated with supporting ISIS [12]. This makes the community of ISIS supporters potentially much bigger than the original set of seed ISIS accounts.

Similar to retweets, if we filter out mentions that were generated by eventually suspended users, we find that ISIS gets more negatively affected compared to a random sample of legit accounts (Figure 5) – another signal that ISIS accounts are more likely to attract other malicious users to interact with them. Thus, we next investigate if such accounts could actually be ISIS related as well.

Are ISIS retweeters pro-ISIS? To answer this question, we collect all tweets generated by retweeters of ISIS accounts (i.e. we collect all their tweets whether ISIS-related or not). We first note that those accounts generated 1.4 billion tweets (15% of all Arabic content in 2015!). We then divide this big set of tweets into two groups: (i) tweets from suspended accounts and (ii) tweets from accounts still alive by end of 2016 (Table 2). We then look into the content created by these two groups. By only looking into the top hashtags used by both groups, we are able to reveal an important finding about those retweeters: Table 3 shows the top-10 hashtags for ISIS and their retweeters. We see an overlap between ISIS and their retweeters who eventually got suspended. Furthermore, Magdy et al., found that Twitter accounts using the Arabic phrase for “#Islamic_State” are 93% likely to be supporters of ISIS [12]. We see that the top used hashtag by the suspended retweeters is indeed “#Islamic_State” hinting that those retweeters are most likely supporters of ISIS. The conclusion here is that the seed set of ISIS users that we started with is apparently the tip of the iceberg and a careful analysis of their interactions could lead to a much bigger community of ISIS supporters.

Are ISIS accounts recruited or born that way? As we see in Figure 1, ISIS accounts manage to post more content before getting

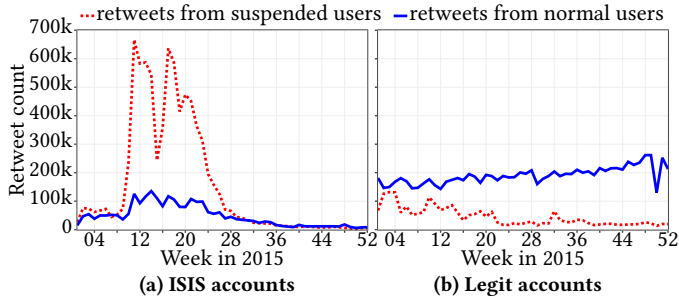


Figure 4: (a) Most (76%) of the retweets of ISIS came from eventually suspended accounts (31% of all retweeters) hinting that a large number of those interacting with ISIS are likely involved in malicious activities as well. (b) By comparison, only 17% of the retweets of legit users were from eventually suspended users (15% of all accounts).

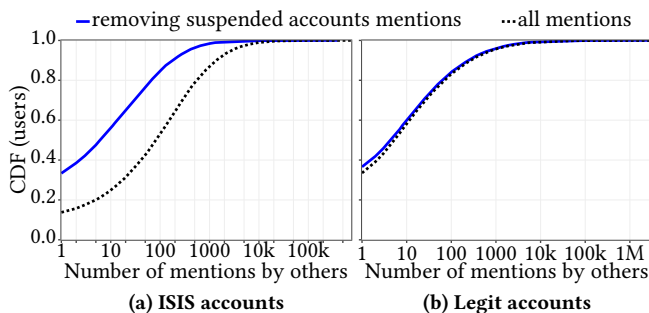


Figure 5: When removing interactions initiated by eventually suspended accounts, ISIS accounts tend to lose a major part of the community interactions (a) while a random sample of legit accounts is negligibly affected (b) – another indicator that a percentage of users interacting with ISIS are also involved in malicious activities.

suspended. Were those accounts *normal* accounts that later turned into bad ones (e.g., as a result of recruitment campaign by ISIS) and hence had more time to spread content? Or were those accounts created to support ISIS from the beginning? A simple way to study this is to check accounts from birth to death. We focus here on accounts that were born in January 2015 and were suspended before December 2015. There are 5,057 such accounts that generated 4,970,042 tweets. We check their top hashtags used at three points in their lifespan: at birth, mid-life, and at death (right before suspension). Table 4 shows that although a majority of accounts have been pro-ISIS from the beginning (because #Islamic_State is the top hashtag all the time), we still see a strengthening of ISIS support over time evident in the increasingly more ISIS related hashtags appearing over users lifespan. For example, we see some innocuous hashtags (e.g., supplications) in the *at birth* column in Table 4 that we don't see at later stages.

5 NEXT STEPS

This initial investigation has shed light on the extent to which extremist groups (ISIS in this study) interact with the online social communities they operate in. Starting from a 24k seed accounts

At birth	Mid-life	At death
Islamic_State	Islamic_State	Islamic_State
Forgiveness	Caliphate_News	Syria
Supplication	Caliphate_State	Caliphate_State
Daesh	Daesh	Nusrah_Front
King_Abdullah_Death	Decisive_Storm	Aleppo
Charlie_Hebdo	Takrit	Daesh
Saudi	Saladin_Area	Fatah_Army
Supplications	Ramadi	Iraq
Forgiveness	Nusrah_Front	Aljazeera
Nusrah_Front	Anbar_Area	Aamaq_Agency

Table 4: Top Hashtags (translated) for ISIS accounts at three intervals in their life: at birth, mid-life, and at death. While the majority of accounts are pro-ISIS from day 1 (#Islamic_State being top hashtag at birth), we see a strengthening message of ISIS overtime, Innocuous hashtags (supplications, forgiveness) disappear at mid-life onward.

known to support ISIS, we were able to uncover a much larger group of potentially ISIS related accounts (170k users). We plan to extend this work by studying the effect of this larger group on the overall Twitter community? e.g., do they function as an *echo chamber* or do they reach out well to the community?

REFERENCES

- [1] Majid Alfifi and James Caverlee. 2017. Badly Evolved? Exploring Long-Surviving Suspicious Users on Twitter. In *International Conference on Social Informatics*. Springer, 218–233.
- [2] Adam Badawy and Emilio Ferrara. 2017. The Rise of Jihadist Propaganda on Social Networks. (2017).
- [3] Matthew Benigni and Kathleen M Carley. 2016. From Tweets to Intelligence: Understanding the Islamic Jihad Supporting Community on Twitter. In *Social, Cultural, and Behavioral Modeling: 9th International Conference, SBP-BRIMS 2016, Washington, DC, USA, June 28–July 1, 2016, Proceedings 9*. Springer, 346–355.
- [4] Jonathon M Berger and Jonathon Morgan. 2015. The ISIS Twitter Census: Defining and describing the population of ISIS supporters on Twitter. *The Brookings Project on US Relations with the Islamic World* 3, 20 (2015), 4–1.
- [5] Elizabeth Bodine-Baron, Todd C Helmus, Madeline Magnuson, and Zev Winkelman. 2016. *Examining ISIS Support and Opposition Networks on Twitter*. Technical Report. RAND Corporation Santa Monica United States.
- [6] Clayton Allen Davis, Onur Varol, Emilio Ferrara, Alessandro Flammini, and Filippo Menczer. 2016. BotOrNot: A system to evaluate social bots. In *Proceedings of the 25th International Conference Companion on World Wide Web*. International World Wide Web Conferences Steering Committee, 273–274.
- [7] James P Farwell. 2014. The media strategy of ISIS. *Survival* 56, 6 (2014), 49–55.
- [8] Emilio Ferrara, Wen-Qiang Wang, Onur Varol, Alessandro Flammini, and Aram Galstyan. 2016. Predicting online extremism, content adopters, and interaction reciprocity. In *International Conference on Social Informatics*. Springer, 22–39.
- [9] Scott Gates and Sukanya Podder. 2015. Social media, recruitment, allegiance and the Islamic State. *Perspectives on Terrorism* 9, 4 (2015).
- [10] Chris Grier, Kurt Thomas, Vern Paxson, and Michael Zhang. 2010. @ spam: the underground on 140 characters or less. In *Proceedings of the 17th ACM conference on Computer and communications security*. ACM, 27–37.
- [11] Gilad Lotan, Erhardt Graeff, Mike Ananny, Devin Gaffney, Ian Pearce, et al. 2011. The Arab Spring! the revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. *International journal of communication* 5 (2011), 31.
- [12] Walid Magdy, Kareem Darwish, and Ingmar Weber. 2015. # FailedRevolutions: Using Twitter to study the antecedents of ISIS support. *arXiv preprint arXiv:1503.02401* (2015).
- [13] Government Publishing Office. 2016. ISIS Online: Countering Terrorist Radicalization and Recruitment on the Internet and Social Media. (2016). <https://www.gpo.gov/fdsys/pkg/CHRG-114shrg22476/content-detail.html>
- [14] Kate Starbird and Leysia Palen. 2012. (How) will the revolution be retweeted?: information diffusion and the 2011 Egyptian uprising. In *Proceedings of the acm 2012 conference on computer supported cooperative work*. ACM, 7–16.
- [15] Khuram Zaman. 2016. Dataset Spotlight: How ISIS Uses Twitter, <https://www.kaggle.com/datasets>, Last Access: 24/11/2017.