Resurgent Insurgents: Quantitative Research Into Jihadists Who Get Suspended but Return on Twitter

[1]Shaun Wright, [1]David Denney, [1]Alasdair Pinkerton, [1]Vincent A.A. Jansen, [1]John Bryden

[1]Royal Holloway University of London

Keywords: ISIS, Twitter, social media, suspension, resurgents, big data, modern terrorism, whack-a-mole

Date submitted 19 October 2015
Accept Submission 07 March 2016
Published on [In Press]

Abstract

Jihadists are very active on Twitter but their accounts frequently get suspended. A debate over the effectiveness of suspension has arisen; an important factor is that Jihadists quickly create new accounts, resurging back like the moles in the "whack-a-mole" game. This causes biases for terrorism and intelligence analysts. Whilst widely acknowledged, little research investigates the problem. In this study we identify resurging Jihadist accounts with novel methods, and provide detailed analysis going beyond previous case-studies. We show that suspension is less disruptive to terrorists than previously thought, whilst the bias and disruption caused to terrorism research has been underestimated.

Introduction

Jihadists have taken to social media. Twitter has emerged as their "favourite" site (Weimann, 2014) and an estimated 46,000-90,000 ISIS supporting accounts were active there in Autumn 2014 (Berger and Morgan, 2015). Jihadists use Twitter for a variety of reasons. The first reason is to spread their messages to a wide audience. The second is for recruitment; the third is to indoctrinate further those drawn to them, like a crucible of radicalisation. And finally, (although not comprehensively) they also use Twitter for seemingly mundane conversation amongst friends.

As a consequence of the volume of data, and its open-source nature, analysis of this source of intelligence about terrorist and extremist activity is becoming more common amongst academics, journalists and government practitioners (Chatfield, 2015; Greene, 2015; Magdy, 2015; Mahmood, 2012; Moriarty, 2015; Ryan, 2014; Stern and Berger, 2015). Whilst there is very detailed research on the Twitter structure and strategies of the top-down, officially-controlled tiers of Jihadist terrorist groups (Stern and Berger, 2015), we argue that the field could benefit from more sustained research on the larger, bottom-up community of Jihadist massed ranks.

Another consequence of how numerous and vocal Jihadists are on Twitter, is the political, cultural and media pressure to take down – or suspend – terrorism supporting accounts (Levy, 2014; Moriarty, 2015). In recent years this has led to several changes in Twitter's suspension policy, and an enormous increase in the number of suspensions. A debate has now arisen in the media and academic literature on the effectiveness of these suspensions (Arthur, 2014; Fisher, 2015; Gladstone, 2015; Stern and Berger, 2015). The assumption is that suspending terrorist supporting accounts reduces the number of terrorists on Twitter. It is assumed that this, in turn, will help counter the objectives for which Jihadists are using social media in the first place: recruitment, radicalisation, spreading propaganda and threats. On the other side of the

debate are concerns over loss of intelligence, freedom of speech, and how realistically achievable the number of suspensions needed to make a dent in the problem is.

Central to this debate is another significant and problematic phenomenon associated with Jihadist social media research: “many of those suspended users simply sat down at their computers the very next day, created new accounts, and started all over again” (Stern and Berger, 2015). This phenomenon is acknowledged in a range of studies (Chatfield, 2015; Magdy, 2015; Berger and Morgan, 2015) and widely referred to as “whack-a-mole” (Arthur, 2014; Berger and Morgan, 2015; Levy, 2014; Stern and Berger, 2015). Those who create these resurging whack-a-mole accounts we call “resurgents” and we provide a more detailed definition later in the paper.

Resurgents do not just cause a whack-a-mole challenge for those performing the suspensions. Their quantity makes identification difficult and so they often go unnoticed. The impact of researchers being unable to identify or control for resurgents is that their datasets will suffer biases; the main bias being replicate error. If the dataset contains duplicate resurgent accounts who get treated as independent data points, this clearly causes errors in any research addressing a range of issues: the number of Jihadist accounts, the level of support for a particular course of action, how unusual a particular behaviour is, and so on.

An example of a problem caused by resurgents is Berger and Morgan's estimate of the (carefully worded) number of “ISIS-supporting Twitter accounts”. The problem is that we do not know how many unique ISIS supporters are represented by these accounts. In another example, Twitter claimed that it had suspended 10,000 ISIS linked accounts in a single day (Gladstone, 2015). Again, it is unknown how many ISIS supporters this represents. These problems occur because there are no methods to identify resurgent accounts amongst this volume of data, or control for the biases that they cause. One of our aims is to help develop such methods and provide these estimates.

It is clear that resurgents cause problems for suspension and for research, yet academic study of them is lacking. Previous studies have discussed suspension and resurgence as a potential flaw with the generalisability of their findings (Berger and Morgan, 2015; Chatfield, 2015; Magdy, 2015). However, almost no research has been done to characterise and describe suspended or resurgent accounts – partly due to the lack of methods for finding them. The impact of resurgents on the effectiveness debate, therefore, currently rests on Stern and Berger's (2015) case study of a single resurgent.

Stern and Berger (2015) conducted a case study of the suspension and single resurgence of the official al Shabaab Twitter account in January 2013 and concluded that suspension is disruptive to terrorists but not to research or intelligence gathering. One of their claims is that finding matching resurgent accounts, and analysing them as continuations of the same account is easy. Furthermore, they claimed the “suspension had cost nothing in intelligence value... and the new account continued the stream of press releases”. Whilst this may be true for researchers tracking a particular case study account, especially official media accounts, any researcher analysing the Jihadist massed ranks on Twitter is going to struggle. We suggest that trying to identify all corresponding resurgent accounts in a dataset of 46,000-90,000 accounts is so time-consuming for humans that there *is* likely to be an intelligence cost. Addressing this hypothesis is another one of our aims in this paper.

Stern and Berger also determined the rate at which their resurgent case-study account accrued followers and calculated that it would take months or years to regain all their followers. They then argued that suspension imposes “clear numeric costs” since ISIS supporters must “reconstruct their social networks and reestablish trust” (Stern and Berger, 2015). While there may be costs for some suspended accounts, this picture is incomplete. We hypothesise that

because Jihadist accounts have previously (and repeatedly) built their reputation and the trust of the community, when they return as resurgents, the nature of Twitter means that they can quickly seek out close comrades from their previous network, initiate contact and re-establish their credentials. Therefore we predict that the number of followers of resurgent accounts should grow faster than naturally growing Jihadist accounts who must establish credentials from scratch rather than simply renew them.

We will also consider other factors that could explain any accelerated growth amongst resurgents. One relevant Twitter phenomenon could be "Follow Friday" (Leavitt, 2014), where participating users recommend accounts (on Fridays) to their followers. These tweets are often signposted with the hashtags "#ff" or "#followfriday", e.g. *"#ff #followfriday @randomuser1 @twitteruser123"*. We hypothesise that they could be helping to drive growth, and will perform an initial test of how common they are to assess the viability of this.

We think that the phenomenon of accounts resurging from suspension is a significant enough feature of modern terrorism to merit further study and definition. With currently only a single case study, we suggest that the next logical step is to study more resurgents, and this is the main aim of our paper. However, since the world of modern terrorist activity is one of social media and big data, conclusions drawn about case studies cannot be appropriately generalised to the whole population of Jihadists. We therefore, as has been identified as necessary in the study of Twitter Jihadists in general, propose using big data methods (Berger and Morgan, 2015) on a large sample of resurgents.

We define a Twitter resurgent as any user who has created multiple accounts on Twitter under different handles (unique user-names beginning '@'). Resurgence does not only occur as the direct result of suspension; some users pre-empt their suspension by changing their handle or operating multiple backup accounts. All resurgent types are included in the definition, however, as they cause the same biases to research datasets. On the other hand we do exclude those who are consciously masquerading as different people (e.g. operating multiple personas or a variety of automatic bots) and we consider the implications of this in the discussion.

In this paper we aim to find sets of accounts belonging to the same resurgents. Once we have done that, we can study and describe them. We will assess how disrupted they are by quantitatively analysing the rate at which they accrue followers compared to non-resurgent accounts, as well as looking at Follow-Friday as a possible driving mechanism. We will also provide an estimate of the proportion of Jihadist accounts which are just duplicates and the proportion which represent unique Jihadists. These findings will give terrorism researchers a better understanding of the true numbers and distribution of Jihadists on social media, as well as an appreciation of how disruptive suspension is for research. We therefore set out the first large scale description of resurgent Jihadists, a significant phenomenon in modern terrorism, challenging, in the process, some of the conclusions about Jihadist social media behaviour drawn by others.

Methods

1. Dataset

The sampling algorithm used was developed to bias sampling toward accounts that tended to have numerous links to other accounts that had already been sampled. The reason for doing this was the principle of homophily: the tendency of people to associate with others similar to them (McPherson, 2001). This principle has been shown to lead to highly intra-linked communities on Twitter that bias their interactions to other members of the community and share a social identity (Bryden, 2011; 2013; Tamburrini, 2015). Consequently, we reasoned that Jihadists would bias the accounts that they followed towards other Jihadist accounts and set up our sampling

algorithm accordingly.
We therefore used weighted snowball sampling (Goodman, 1961) to identify Jihadist Twitter accounts. This approach enabled us to grow the sample, whilst weighting sampling towards accounts with numerous links to accounts already identified. A handful of publicly-known, official “media” Jihadist Twitter accounts named by newspapers provided our starting point. We then manually inspected the Twitter followers of these accounts, aided by Twitter's “Who to follow” algorithm, and from our analysis we identified 34 'unofficial-but-supporting', Jihadist accounts. For practicality, we selected only English speaking accounts. We then used this starting sample to seed the snowball algorithm.

We snowball sampled daily between May and July 2015 (77 days, with power issues preventing sampling on 10 days). On each day we looked at all accounts followed by those already in our sample. We then sampled any account identified as being followed by >10% of the users in our sample, and with <1,000 followers of its own.

We selected the 10% threshold to grow the sample slowly, without accelerating, whilst remaining within a relatively tight community of English speaking Jihadists (the principle of homophily). While our sample was smaller than 100 users we used a fixed threshold (new accounts must be followed by more than 10 accounts in our sample). We switched to the 10% threshold once we had sampled 100 accounts.

The upper limit of 1,000 followers was selected for two reasons. Firstly, to prevent the inclusion of popular journalists and academics who are often both highly interlinked with the networks, and connected outwards to non-Jihadist followers. Such community transcending journalists were liable to divert the sampling away from the Jihadist community. Our cut-off is similar to, although more ruthless than, the precedent set by Berger and Morgan (2015) who used a 50,000 cut-off, finding that accounts more popular than this were unrelated. Secondly, by avoiding the more 'popular' accounts, we aimed to direct our dataset away from the official, top-down Jihadist media accounts covered in other research, and towards the largely neglected Jihadist massed ranks.

During sampling, some accounts were protected, suspended or had voluntarily changed their user-name. We moved these to an “inactive sample” where we recorded all the account information, but they no longer contributed to the 10% threshold check. We identified suspended users by the official suspension report with which Twitter had replaced their pages. Protected users had activated privacy settings and only biography, pictures and summary meta-data were available. Non-existent accounts display an official Twitter message that the user cannot be found (despite our evidence that they previously did). Although no information is provided about their non-existence, since Twitter does not report them as suspended we assume that the users changed their handles themselves.

2. Finding resurgent Jihadists

To identify resurgent accounts we used a quantitative approach that helped draw our attention to accounts whose Twitter biographies, names and locations contained at least 30% of the same words We set out the rationale for why our quantitative approach in needed, over human identification of accounts, in Supplementary Material 1. We then visually assessed those accounts, identifying and classifying resurgents.

Defining a set of accounts belonging to a resurgent

When comparing accounts, we used open criteria for determining whether they formed a set. However, in practically all cases, an almost identical match between handle, name, biography or location was necessary and sufficient. Biography and handle were the strongest indicators, whilst location, surprisingly, was still informative due to peoples' unique spelling, punctuation, and choice of descriptive terms. A hypothetical, illustrative example of an almost identical match would be the handles “@jihad_bob2” and “@jihad_bob3”.

Figure 1. An illustrative example from our data of two resurgent accounts which we classified as a set. They have almost identical handles and almost identical biographies. Their images were not inspected, but their profile images are an almost identical match too. Screenshots of two user accounts taken from http://twitter.com.

Figure 2. An illustrative example from our data of two accounts which we did not classify as duplicates of one another, despite some similarities. Screenshots of two user accounts taken from http://twitter.com.

3. *Do resurgents accrue followers faster?*

We investigated how disrupted resurgent accounts are by calculating their rate of follower accrual versus non-resurgent controls. As we were unable to find other matching resurgents, we treated all those who had not been identified as non-resurgent controls. We calculated growth rate by dividing the number of followers an account had upon sampling by the number of days between creation and sampling. We used the non-parametric Mann-Whitney U test after ruling out normality (both p = 0.00, 2.d.p, Kolmogorov-Smirnov).

Results

We sampled 1,920 English speaking Jihadist accounts from Twitter. By the end of sampling 1,080 had been suspended, 141 accounts were private, 97 no longer existed and 602 were active (figure 3). Only 1,858 of the users had sufficient

name, location and biography information for analysis.

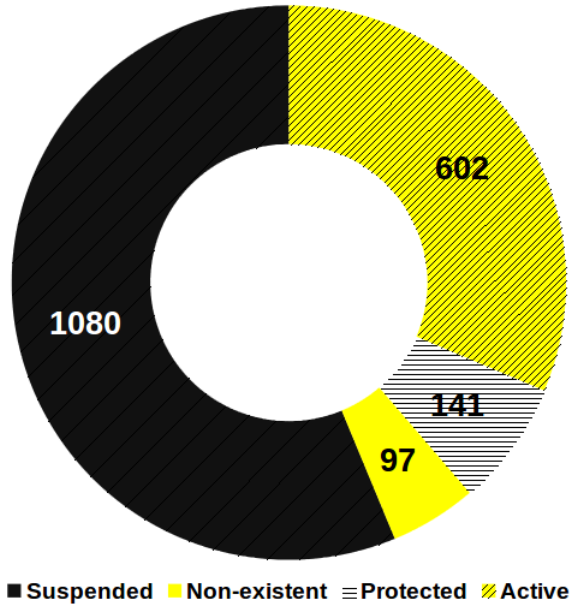

Figure 3. The distribution of our dataset of 1,920 English speaking Jihadist Twitter accounts. By the end of sampling, 1,080 had been suspended by Twitter, 141 had set their accounts to private, 97 no longer existed due to voluntary name change and 602 were still active.

1. *Terrorist group affiliations*

The majority of accounts do not declare a terrorist organisation affiliation, nor does a simple content analysis allow for unequivocal categorisation. Amongst 300 randomly selected users, 39 (13%) provided an allegiance, of which all gave ISIS, IS, Islamic Caliphate, Baqiya or Khilifa. Amongst the 261 that didn't, 34 (13%) gave one of the four most common locations: "Dar ul Kufr" [Land of the unbelievers] (n=16, 6%), "UK" (n=12, 5%), "Dunya" [the non-spiritual, temporal world] (n=3, 1%) , and "Somalia" (n=3, 1%); with the sharing of extremist content and pro-Caliphate sentiment also common. Twitter also suspended 56.3% of our sample, evidence that suggests they were engaging in extremist activity. We therefore categorise our sample as Jihadists, whilst assuming, based on location and content, that the majority are ISIS-supporting members of the "Baqiya family" (Amarasingam, 2015).

2. *Finding resurgent Jihadists*

Using the quantitative approach outlined in the methods, we estimated the number of unique Jihadist users by identifying resurgents: users in the dataset who had multiple, matching replicate accounts.

From 1,858 user accounts with information to analyse, only 1,484 (79.9%) were unique Jihadists. The remainder, over one in five accounts, were duplicates: resurgent accounts. 192 (12.9%) of the unique users were resurgents who owned, on average, 2.95 accounts (a set of mean size 2.95) within the three month period (table 1).

The other statistic commonly reported is the number of Jihadist accounts that have been taken down or suspended. This also overestimates the number of unique Jihadists. Performing the same analysis with the suspended users with information to analyse (n=1,066), we found only 757 (71.0%) unique Jihadists. 114 (10.7%) of these unique users were resurgents, owning a mean of 3.71 suspended resurgent accounts in three months (table 1).

	Entire Sample	**Suspended Users**
Accounts analysed	1,858	1,066
Unique Jihadists	1,484 *(79.9%)*	757 *(71.0%)*
Duplicate accounts	374 *(20.1%)*	309 *(29.0%)*
Unique users who were resurgents	192 *(12.9%)*	114 *(10.7%)*
Mean # of accounts belonging to each resurgent	2.95	3.71

Table 1. Identification and quantification of resurgents in the dataset: users who had multiple, matching replicate accounts.

3. *Resurgents accrue followers faster*

We found that the growth rate of resurgent accounts (n=566, median 43.8) is significantly greater ($p < 0.0001$ [exact p-value $< 2.38 \times 10^{-40}$], 1-tailed Mann-Whitney U) than that of naturally growing, non-resurgent accounts (n=1,292, median 8.37) (figure 4).

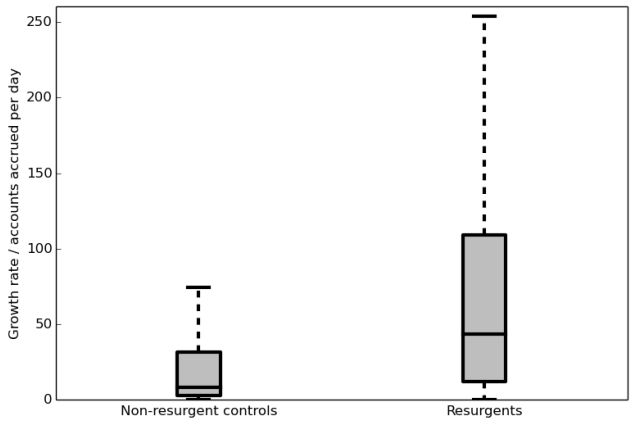

Figure 4. The growth rates (followers accrued per day) of resurgent accounts (n=566) versus naturally growing Jihadist accounts (n=1,292). The plot shows the growth rate of resurgent accounts is significantly higher than that of non-resurgent accounts.

4. *Jihadist Follow-Friday*

Having shown that resurgent accounts grow faster than those of non-resurgents, we searched for explanatory factors. We observed a similar phenomenon to "Follow-Friday" within the Jihadist Twitter community and assessed their viability as a growth driving mechanism by testing how common these tweets were.

Downloading the entire daily tweet output of our sample generated a corpus of approximately 155,000 tweets. We randomly-selected 2,500 tweets from this corpus; 46 (1.84%) fitted the Jihadist Follow-Friday structure.

Although we dub them "Jihadist Follow-Friday" tweets, zero (0.0%) contained Friday hashtags. Furthermore, none (0.0%) of the 46 examples promoted more than one user per tweet, with 17 (37.0%) repeating the name several times per tweet, e.g. "*Follow: @jihadistaccount123 @jihadistaccount123 @jihadistaccount123"*, and the remaining 29 (63.0%) naming them only once, e.g. "*FOLLOW & SUPPORT @jihadistaccount123*". Three tweets (6.52%) also stated that the user had returned from suspension.

As an indicator of whether Jihadist Follow-Friday tweets are significant enough to contribute to re-growth, this result estimates that there are 2,852 tweets (1.84%) promoting other Jihadist accounts in our dataset of 155,000 tweets.

Discussion

Suspension and resurgence are significant phenomena in modern, online terrorism. As resurgents are difficult to find in large numbers, research into them is scarce, relying on Stern and Berger's (2015) case study alone. Furthermore, terrorism research treats the duplicate resurgents as independent data points, biasing social media research into the numbers, opinions and behaviour of Jihadists. We found resurgents, estimating that within our sample only 79.9% of Jihadist Twitter accounts belong to unique Jihadists, with a lower 71.0% of unique Jihadists amongst suspended accounts. This gives researchers a better picture of the patterns displayed by resurgents, as well as a scale of the significant biases for research and estimates and the continuous disruption to intelligence gathering.

With the identification of resurgents comes the ability to analyse them beyond individual case studies. Previous work has concluded that there are "clear numeric costs" to resurgents who suffer slow regrowth as a cost of suspension (Stern and Berger, 2015), contrary to this single al-Shabaab account however, we have shown that in our sample resurgents grow significantly faster (median 43.8 accounts accrued per day) than non-resurgent Jihadists (median 8.37). Whilst it remains possible that this might not be sustained long enough to get back all of their old followers, especially the curious Westerners,

there is no obvious disruption to Twitter when considered as a crucible of radicalisation. Whether or not Jihadist Follow-Friday tweets help to drive this accelerated growth also merits further study, as they seem prominent (1.84% of tweets) given the number of alternative discussion topics.

Our findings could help analysts to put reported numbers and statistics in a more appropriate context. Berger and Morgan estimated the number of ISIS supporting Twitter accounts at 46,000-90,000. However, we have shown that an improved estimate should drop below 36,800-72,000 (79.9%) unique users. Another commonly reported, headline-catching statistic is the number of ISIS accounts suspended; Twitter reported suspending 10,000 accounts. However, our results suggest that this should be corrected to represent only 7,100 (71.0%) unique ISIS supporters. We suggest that while the rate of suspensions remains stable, our specific results of 79.9% (overall) and 71.0% (amongst suspended) may have some usefulness, but that even when suspensions escalate, the principle behind our finding remains crucial. All of these results highlight the dangers in working with a Jihadist dataset without correcting it for resurgents.

One of the implications of this improved picture of resurgents is the contribution to the suspension effectiveness debate. A great deal of political and public pressure exists to suspend terrorists and their supporters from social media sites. Although intelligence concerns often take "a distant third" place to business and cultural concerns, some argue that the intelligence costs themselves are limited (Stern and Berger, 2015). Whilst our results do not address the cultural or ethical arguments, they do suggest that suspensions are less disruptive to terrorists than previously argued; furthermore, suspensions cause significant biases to data and its analysis. Rather than leading us, however, to advocate against suspension – there are convincing ethical and intelligence quality improving arguments (Stern and Berger, 2015) – we propose using methods to control for it.

We consider our dataset of accounts, and their suspension rates, to be generalisable to the unofficial, English-speaking, Jihadist community on Twitter. We categorised our sample as pro-ISIS members of the "Baqiya family" (the friendly network of online ISIS supporters) (Amarasingam, 2015), although terrorist group affiliation is almost impossible to assess without additional sources of data. It is, however, in line with the political dominance of ISIS during summer 2015, the nature of the "Baqiya family" (Amarasingam, 2015), and Berger and Morgan's (2015) estimate of 46,000-90,000 ISIS-supporting accounts during a similar length time. Although it is possible that generalisability is limited by snowball sampling's bias towards the seed list, after sampling 1,920 accounts from a seed list of 34, any initial bias should have been diluted. We therefore associate our results only with the general "Jihadist" community, limiting the ability of our study to make statements about differences between specific terrorist groups. Inspection of the data does, however, indicate success in our aim of using a minimum popularity to exclude bots.

A potential critique of our sampling method (continually looking for new accounts) is that it could be biased towards resurgents. We defend the appropriateness of our sampling, however, as it will still snowball into a wider community, reaching out to newly discovered accounts that need not be new to Twitter. We would also point out that although snowball sampling cannot reach disjoint groups, such a hypothetical, unconnected terrorist group is by definition unrepresentative of the ISIS-dominated Twitter environment. We do, however, suggest that the best course of action is for researchers themselves to analyse their dataset for resurgents. Finally, our definition of resurgents also excludes those masquerading as bots or multiple personas. These are phenomenon potentially causing additional replicate biases to terrorism research and therefore merit further research.

Although there appear to be some statistical issues with generalising our findings directly to

Berger and Morgan's work, there are several possible counter-explanations. Scaling by 79.9% predicts that over 20% of their users have resurged back, but they only reported ~7.5% being suspended in the first place. There are however, three reasons why this need not contradict our findings, nor stop us applying our result to their data. Firstly, they acknowledge that the suspension rate has dramatically escalated since, and in our data it was 56.3%. Secondly, name-changing and backup accounts are also sources of resurgents and are presumably not covered under their reported suspension statistics. Finally, it appears that their sample was not continuously re-checked for suspensions. Thus their suspension rate may actually be higher than reported. In the specific case of our Twitter example, where all the accounts were active during a single day, our findings may also not be applicable. However, whenever accounts are reported suspended over a period greater than several weeks, our findings may be highly informative. Again, these challenges only emphasise the importance of researchers attempting to find resurgents in their data for themselves.

Our study included several types of resurgents, including backup accounts and those created after suspension. The difference between a backup and post-suspension account is not a binary classification but a spectrum, depending on whether the main account has been suspended, the age of the backup before and since becoming the main account, and the ratio between these. Recording data to investigate these is therefore beyond the scope of this article, but merits a future study. Crucially however, a lower rate for backups would lower the rate for combined resurgents, and this thus indicates the robustness of our significantly elevated result.

Limitations

A limitation of our “Jihadist” study is that we cannot make statements about the differences between specific terrorist groups. These findings could also benefit from more work with a broader sampling procedure, as there are limits on generalising our sample to the unofficial, English-speaking, Jihadist, Twitter community (snowball sampling methods both limit the ability to reach disjoint groups, and exhibit bias towards their seed lists). Additionally, our estimates are conservative upper bounds as we could have missed some resurgents due to the challenge of finding resurgents amongst big data. Our estimates are also upper bounds as our definition excluded those masquerading as bots or multiple personas, and our study amalgamated several types of resurgents, including backup accounts and those created after suspension. Although there are likely to be differences between backup and post-suspension resurgent accounts (we hypothesise that their longer lifespan and insignificance to followers would give backup resurgents a lower growth rate), testing this is non-trivial. There may also be limitations with generalising our findings directly to all other numerical estimates, as sampling methods differ from study to study.

Conclusion

This paper marks a step change in methodological approaches towards the study of resurgent Jihadists. The new methods give us novel insights into the proportion of fast-growing, duplicate accounts (20-30%), which in turn suggest some crucial new approaches in terrorism studies: adjusting numerical estimates, recognising dataset biases, and seeking methods to identify and control for the significant number of resurgents. Our quantitative method in particular, which we hope to calibrate further in future work, appeared very useful for quickly finding resurgents, and this presents a clear example of the wider importance and power of using quantitative analysis to investigate a range of terrorism behaviours.

Author contributions

All authors were involved in the conception of the work. SW collected the data, performed the analyses, wrote the first draft and led the writing of the manuscript. DD, AP, VAAJ and JB edited and critiqued the manuscript. The authors would also like to express their gratitude to Peter Adey for helpful discussions and feedback on the manuscript, and to the reviewers for their constructive comments.

About the authors

Shaun Wright *is a PhD candidate at Royal Holloway University of London where he is also a visiting teaching assistant and guest lectures on the Terrorism, Insecurity and Risk course. Reflecting the interdisciplinary nature of terrorism studies, he has supervisors in the departments of Biological Sciences, Law, Criminology & Sociology, and Geography. He studied for his undergraduate BA at Gonville & Caius College, Cambridge University, combining courses in Computer Science, Psychology and Neuroscience.*

David Denney *is Professor of Social and Public Policy in the School of Law, Royal Holloway University of London. He has written extensively on various aspects of risk in society, violent crime and discrimination in the criminal justice system. He is currently Principal Investigator on ESRC Dstl funded research which examines the impact of social media on military personnel and their families.*

Alasdair Pinkerton *is a political geographer with particular interests in issues related to critical geopolitics, the media, and the 'international relations' of public diplomacy. He has regional specialisms in the South Atlantic and South Asia, as well as working extensively in the US, Canada, the UK and Cyprus. He appears frequently in the UK and international media on issues related to global geopolitics.*

Vincent Jansen *is a Professor of Mathematical Biology at Royal Holloway University of London. He is interested in population dynamics and evolution, including the dynamics and evolution of social systems. His research entails the formulation and analysis of mathematical models to understand biological processes, and he has applied this to a range topics in biology and beyond. Part of his research has addressed how social networks form, how people communicate on social networks and how our language is shaped and influenced by the social network that we are embedded in.*

John Bryden *is a Research Fellow at Royal Holloway University of London interested in group behaviour, dynamics and evolution in biological systems and human societies. More generally, he is interested in the modelling of complex systems and the methodological issues raised by the modelling of complex systems. He is building a methodological approach for the studying of human behaviour in social systems by applying models to large data sets. He has recently published work studying how humans cluster into groups on social networks and the effects this has on their language.*

References

Amarasingam, A., 2015. What Twitter Really Means For Islamic State Supporters. *War on the Rocks*. December 30.

Arthur, C., 2014. Taking down Isis material from Twitter or YouTube not as clear cut as it seems. *The Guardian*. June 23.

Berger, J.M., Morgan, J., 2015. The ISIS Twitter Census: Defining and describing the population of ISIS supporters on Twitter. *The Brookings Institution.*

Bryden, J., Funk, S., Geard, N., Bullock, S., Jansen, V.A.A., 2011. Stability in flux: community structure in dynamic networks. *Journal of the Royal Society Interface*. 8(60):1031-40.

Bryden, J., Funk, S., Jansen, V.A.A., 2013. Word usage mirrors community structure in the online social network Twitter. *EPJ Data Science*. 2(3)

Chatfield, A.T., Reddick, C.G., Brajawidagda, U., 2015. Tweeting propaganda, radicalization and recruitment: Islamic state supporters multi-sided twitter networks. *Proc 16th Annual International Conference on Digital Government Research.* pp.239-49 DOI:10.1145/2757401.2757408

Fisher, A., 2015. Swarmcast: How Jihadist Networks Maintain a Persistent Online Presence. *Perspectives on terrorism*. 9(3)

Gladstone, R., 2015. Twitter Says It Suspended 10,000 ISIS-Linked Accounts in One Day. *The New York Times*. April 9.

Goodman, L.A., 1961. Snowball Sampling. Ann Math Stat 32:148–170.

Greene, K.J., 2015. ISIS: Trends in Terrorist Media and Propaganda. *International Studies Capstone Research Papers.* Paper 3.

Leavitt, A., 2014. From #FollowFriday to YOLO: Exploring the Cultural Salience of Twitter Memes. In: Weller, Bruns, Burgess, Mahrt, Puschmann. ed. *Twitter and Society*. New York: Peter Lang. pp.137-154.

Levy, R., 2014. ISIS Tries to Outwit Social Networks. *Vocativ*. June 17

Magdy, W., Darwish, K., Weber, I., 2015. #FailedRevolutions: Using Twitter to Study the Antecedents of ISIS Support. arXiv:1503.02401v1

Mahmood, S., 2012. Online social networks: The overt and covert communication channels for terrorists and beyond. *2012 IEEE Conference on Technologies for Homeland Security (HST).*

McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*. 27:415-44.

Miller, G.A., 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*. 63(2):81–97.

Moriarty, B., 2015. Defeating ISIS on Twitter. *Technology Science*. 2015092904

Ryan, L., 2014. Al-Qaida and ISIS Use Twitter Differently. Here's How and Why. *National Journal*. October 9.

Standing, L., 1973. Learning 10,000 pictures. *Q J Exp Psychol.* 25:207-22.

Stern, J., Berger, J.M., 2015. *ISIS: The state of terror.* London: William Collins

Tamburrini, N., Cinnirella, M., Jansen, V.A.A., Bryden, J., 2015. Twitter users change word usage according to conversation-partner social identity. *Social Networks.* 40:84-89.

Weimann, G., 2014. *New Terrorism and New Media*. Washington, DC: Commons Lab of the Woodrow Wilson International Center for Scholars.

Supplementary Material 1

Why a quantitative approach?

Manually inspecting the complete dataset of 1,920 users for replicates would be very time consuming. Berger's sample of 46,000+ would make the task close to impossible. The feasibility of this task is partly limited by its reliance on human memory capacity. Whilst working memory capacity is a mere 7±2 items (Miller, 1956), we suggest that a more appropriate indicator is recognition memory – the ability to recognise whether or not something matching the account had been encountered earlier in the dataset. Standing (1973) empirically derived equations showing that recognition memory follows a power law with the number of items presented. We can therefore calculate that if humans inspected our 1,920 accounts as printed words, Standing's work predicts the number capable of being held in memory is:
10^((0.92*log(1,920 items))-0.01) = 1,025

Since for many accounts we also have a screenshot of their Twitter profile, Standing's equation for pictorially presented data predicts:
10^((0.93*log(1,920 items))+0.08) = 1,360

The upper limit of human memory whilst attempting a match search with our medium sized dataset is therefore ~53-71% of previously encountered accounts. Since each account is actually represented by around 10 words, not one, this oversimplification generates an extremely conservative upper limit. Re-calculating for Berger's conservative estimate of 46,000 ISIS accounts, only ~41-56% can be held in recognition memory; another overestimation. Standing's results may also not generalise this far beyond the 10,000 items used in his work.

To aid the quick finding of resurgent accounts, we therefore used a quantitative approach to draw our attention to several accounts at a time. Hypothetically, the simplest approach might have been selecting two random accounts to evaluate simultaneously. This would have been ineffective. A quantitative approach should work on an assumption or hypothesis about the data. We assumed that finding matches would be aided by selecting accounts whose biographies, names and locations contained >30% of the same words. This meant that only accounts with these attributes had sufficient information to analyse.