



Violent extremist group ecologies under stress

Manuel Cebrian^{1,2,3}, Manuel R. Torres^{4,5}, Ramon Huerta⁶ & James H. Fowler^{7,8}

SUBJECT AREAS:

POPULATION DYNAMICS

PSYCHOLOGY AND BEHAVIOUR

BEHAVIOURAL ECOLOGY

ECOLOGICAL NETWORKS

Received

24 August 2012

Accepted

25 February 2013

Published

27 March 2013

Correspondence and requests for materials should be addressed to

M.C. (manuel.

cebrian@nicta.com.

au)

¹Media Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA, ²Department of Computer Science and Engineering, University of California at San Diego, La Jolla, California 92093, USA, ³NICTA, University of Melbourne, Melbourne, Victoria 3010, Australia, ⁴Political Science Department, Pablo de Olavide University, Seville 41014, Spain, ⁵Weatherhead Center for International Affairs, Harvard University, Cambridge, Massachusetts 02138, USA, ⁶Biocircuits Institute, University of California at San Diego, La Jolla, California 92093, USA, ⁷Medical Genetics Division, University of California at San Diego, La Jolla, California 92093, USA, ⁸Political Science Department, University of California at San Diego, La Jolla, California 92093, USA.

Violent extremist groups are currently making intensive use of Internet fora for recruitment to terrorism. These fora are under constant scrutiny by security agencies, private vigilante groups, and hackers, who sometimes shut them down with cybernetic attacks. However, there is a lack of experimental and formal understanding of the recruitment dynamics of online extremist fora and the effect of strategies to control them. Here, we utilize data on ten extremist fora that we collected for four years to develop a data-driven mathematical model that is the first attempt to measure whether (and how) these external attacks induce extremist fora to self-regulate. The results suggest that an increase in the number of groups targeted for attack causes an exponential increase in the cost of enforcement and an exponential decrease in its effectiveness. Thus, a policy to occasionally attack large groups can be very efficient for limiting violent output from these fora.

Emerging technologies that permit data collection at unprecedented scale can reveal patterns of individual and group behavior^{1–5} and they have propelled the development of data-driven mathematical models of human conflict. These include models of insurgent and terrorist activity^{6–10}, global trends in violence across human history^{11–13}, cultural and ethnic confrontation^{14–17}, and the structure and dynamics of criminal organizations^{18–20}.

Here, we turn our attention to violent extremist groups that make intensive use of Internet fora for recruitment, illegal financing, internal communication and, in fully radicalized individuals, operative support of real world violent activity^{21–24}. For the past decade, extremists have been using online networks to socialize their members and to reinforce aspects of their radical culture^{25,26}. These networks diminish the sense of isolation for members of extremist groups and give them access to a global community that enjoys the legal and religious approval of the leaders of their movements²⁷. They also facilitate the acquisition of articles, books, and audiovisual material with radical content that are created in clandestine or underground circles. Without access to these networks, obtaining material like these would be much more difficult²⁸.

Online fora are particularly important for extremist groups, as they play a key role in the terrorist recruitment process. These platforms have not only become a common medium for extremist organizations to spread their propaganda and obtain funding, but they have also transformed into the primary arena where the process of violent radicalization takes place²⁹. Extremist fora allow users to reinforce their identities as militants through virtual activism and help to create a hierarchy within the group that can be easily monitored by the group leaders. In fact, it is possible to detect specific participation rules in these fora that allow users to improve their reputation within the group and their status within the community³⁰. Those users who achieve the highest status may then be recruited to engage in real world operations^{21–24}.

Although it is well documented that recruitment into online fora can lead to recruitment into terrorist organizations, it is important to emphasize that these processes are not the same. We are not claiming that online fora are a proxy for real world terrorism. Rather, we are claiming that they are an important entry point for some people who later engage in real world acts of violent extremism, and it is therefore important to understand how these fora operate and how we might act to reduce their effectiveness.

Several thousand websites support and promote extremist discourse, but not all have the same impact. A very small number of sites exert a very large amount of influence because they are the only ones to directly receive


Table 1 | List of online fora analyzed. Forum names and URLs are available from the authors on request

Forum	Starting recording date	Last recording date	Total number of days recorded
Forum 1	06/07/2008	07/03/2011	40
Forum 2	01/01/2008	02/25/2011	7
Forum 3	02/23/2010	06/23/2011	44
Forum 4	08/02/2008	07/03/2011	46
Forum 5	02/25/2011	07/03/2011	39
Forum 6	02/25/2011	07/03/2011	39
Forum 7	01/17/2009	03/23/2011	21
Forum 8	10/29/2008	07/03/2011	49
Forum 9	02/05/2011	07/03/2011	11

material prepared by extremist groups^{31–36}. On these fora one can find ideological diatribes, interviews, and audio and video recordings of recent acts of violence. The other sites simply repeat and amplify the content disseminated by the most prominent extremist fora.

Extremist Internet fora operate in a hostile environment as their growing importance has caused them to come under scrutiny in recent years. Indeed, the conventional wisdom is that the longevity of an online community is inversely proportional to its importance in the radical universe. Organizations dedicated to monitoring extremist groups on the web note that 80% of the top websites to emerge between 2002 and 2004 have disappeared because they have been unable to withstand continuous harassment by security agencies, private groups, and hackers^{37,38}. These attacks are designed to infiltrate and boycott extremist fora³⁹ and they have also succeeded in provoking mistrust and mutual accusations between rival groups²¹. As a result, administrators are suspicious of new users who are too active or who express a clear interest in contacting forum members in person. Some platforms have implemented stricter self-protection measures such as making their entire content available only to registered users, closing admission to new members, or accepting only new users who have been endorsed by a current one.

Actors seeking to disrupt recruitment to extremist fora face several technical questions^{40–43}. Which fora should be targeted for an attack? What strategy should be used? Given that these attacks require costly human resources to be implemented, it is important to assess the impact of various strategies, and to do this we need data and a model that will help us to better understand how fora react when they are under stress. Elucidating these questions will shed light into an important debate: Should terrorist groups be allowed to use the Internet, or should they be banned from it⁴⁴?

In order to analyze the potential impact of cybernetic attacks, we recorded the number of members registered from 10 different fora, as reported automatically by the forum platforms (see Table 1, Table 2, and Figure 1 for the list of forums analyzed and monitoring statistics). Recordings were performed at 12 AM GMT. The data collection started 1/1/2008 for the Forum 2, which is the earliest forum that we were able to detect, and spanned the next 4 years for the rest of the

fora. The number of measurements per forum ranges from 7 to 49, and the sampling rate was constrained by the availability and accessibility of each forum. In Table 2 we provide the details of the data collected for each forum and overall statistics showing, among other things, the skewness of the distribution (notice that the mean number of members is significantly larger than the median). Forum 10 is not shown here, as it was discarded from our analysis for reasons discussed below.

The reported information was publicly accessible without login or participation required. No further checks on the validity of these reported figures were performed, so we cannot rule out systematic biases introduced by the forum administrators (e.g. multiple registrations by the same users, counting of inactive users, etc.). During the observation period we did not detect any deliberate efforts to strategically misrepresent the data. However, given the possibly quite large incentives for misrepresentation, further work on the assessment of data veracity is certainly needed. We also cannot rule out problems arising from the left-censoring of our data, as we lack information about activity prior to the detection of a forum.

Although the observation period for the 10 fora spans 4 years, we were unable to record information on a daily basis, as we would find the fora inaccessible on many occasions – mostly de-activated with no website connectivity. The time-span for which these fora would be de-activated varied from days to months. Even though we may suspect that the fora under study were shut down as a result of an attack, we do not have direct evidence of this attack, and cannot establish a causal link.

In selecting the sites for analysis, we focused attention on fora that constituted the “inner circle” or core of the violent extremist presence on Internet. These fora 1) served as primary sources of dissemination of propaganda produced directly by extremist organizations, 2) had users that were implicated in real-world acts of terror, and 3) had the largest following and influence in the extremist cyber-community^{31–36}. In addition, we only included a particular forum if we detected at least one instance where sharing of text or multimedia material prepared by extremists took place during the observation period. These fora also contained normal discussion and sharing of non-extremist information among members.

We excluded from our analysis “secondary” fora that hosted radical content copied from other sites because they typically attract very few users. This also lead us to discard one of the ten fora under study, run from Spain prior to August 2011. During its three-month existence, it only managed to attract 13 registered users despite its more than 7, 800 discussion topics and 10, 000 posts from other sites, which were uploaded daily by the site administrator⁴⁵.

In Figure 1 we show the rate of growth for each forum, and given the different initial start dates we also show these curves fitted to reduce the Euclidean distance to Forum 8, which had the most observations. These data suggest that fora exhibit a first phase of exponential growth followed by a saturation phase in which they continue to grow exponentially but at a much slower rate. This contrasts with the constant exponential growth (with probably a very large cutoff established by the size of the susceptible population) one

Table 2 | Monitoring statistics for the fora analyzed. The statistics are calculated for the full observation period and all fora, with the aim to give an intuition on the large variability in our ability to collect data across different fora

Number of Fora	9
Starting recording date	1/1/2008
Final recording date	7/3/2011
Total number of days recorded	296
Min value	13
Max value	14,135
Mean value	5778.1
Median	4912
Standard deviation	4228.1

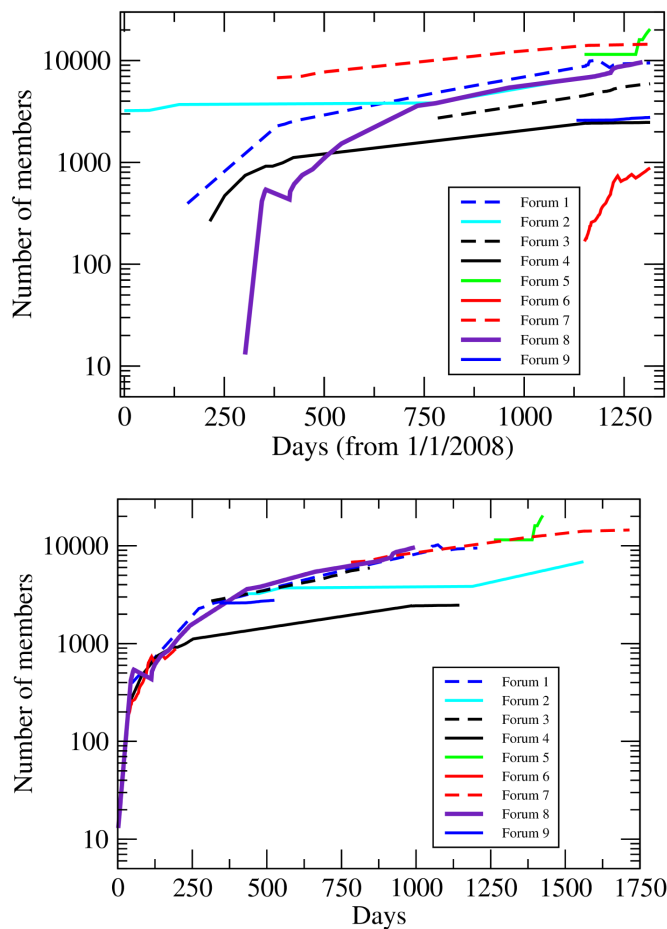


Figure 1 | (TOP) Number of members participating in the 9 extremist fora under study. Note that initially the fora grow exponentially and then saturate growing at an exponential yet much slower rate. (BOTTOM) Fora time shifted to match the early stages of the development of every fora and gain visual intuition on their dynamics. In particular, we aligned the early stages of the all fora with Forum 8, with the exception of Forum 5, which was aligned with Forum 9 due to a similar volume of members.

might expect if members were recruited and allowed to join without restrictions (as in the classic susceptible-infected epidemiological model^{46,47}). Case studies suggest that extremist fora control growth by limiting the number of participants⁴⁸.

In what follows, we describe a dynamic characterization for the evolution of one forum, and then the population dynamics of multiple interacting fora. Our goal is to provide a plausible explanation for the empirical observations and to understand the impact of persistent disruptive attacks on the evolution of these online communities.

Results

We start by denoting N as the total population that is susceptible to becoming a forum member and let $\pi_{1,0}$ be the rate of becoming a forum member. Within the forum itself there are several levels of status. We are interested in the number of users at each level, so let $X_i = X_i(t)$ be the number of active members at level i . We assume that there is a process for advancement that allows users to go from one level to another. The rate of transition between levels is given by the transition probability matrix $\pi_{i,j}$ which denotes the probability per unit of time of moving from level j to level i .

For simplicity, we assume the transition probability matrix can only move upwards, hence $\pi_{i,j} = 0$ for $j \geq i$. It may be the case that

some users exit the process at different points through withdrawal (i.e. de-registration), inertia, lack of Internet access, and so on. However, in our model, the probability of exit is absorbed into the transition probabilities from one level to the next one, where a probability of disengagement would yield a smaller probability of transition upwards. We also assume that the number of law enforcement and intelligence agents inhabiting these fora are much smaller than the number of genuine members, as otherwise it would raise awareness of the forum administrators and compromise their infiltration operation.

Suppose that there are L levels within the forum. At the highest level an individual may become fully recruited to engage in real world operations (O) or to become a recruiter (R). We denote O and R the output of the model and recruiters, respectively, produced by the forum. Recruitment is the bottleneck of the radicalization process, as the number of recruiters determine the change rate of X , R , and O . This is implicit in the evolution of X_L , as in this equation, the members leaving the forum at level L depend on R , not on X_L .

The dynamics of the populations within a particular forum can be written as

$$\begin{aligned} \frac{dX_1}{dt} &= \pi_{1,0}N - \pi_{2,1}X_1, \\ \frac{dX_i}{dt} &= \pi_{i,i-1}X_{i-1} - \pi_{i+1,i}X_i, \\ \frac{dX_L}{dt} &= \pi_{L,L-1}X_{L-1} - p_{O,R}R - p_{R,R}R, \\ \frac{dR}{dt} &= p_{R,R}R, \\ \frac{dO}{dt} &= p_{O,R}R, \end{aligned} \quad (1)$$

where $2 \leq i \leq L-1$, and $p_{R,R}$ and $p_{O,R}$ denote the transition rates to R and O , respectively. In these equations we assume that the number of recruiters is smaller than the number of members because there is a bottleneck in the conversion from being a forum member to being a recruiter (R) or engaging in real life terrorism (O). We also assume that the total population is much larger than the number of users in the forum $N \gg X$.

The model itself is a variation of a mass-action compartment model familiar in mathematical epidemiology^{46,47}. Conceptual simplicity drives the choice of this model, though we must stress the caveat that changes in psychological state or activity may not necessarily be similar to changes in disease status.

Full observation of all the states L , O , R presents ethical, experimental^{49,50}, and legal complications⁵¹, as it requires interaction with the Forum participants for long periods of time. Thus, we present a simplified version of the model where we merge all the levels of the fora into one: $X = X_1 + X_2 + \dots + X_L$. Thus, we can rewrite the previous equations as

$$\begin{aligned} \frac{dX}{dt} &= \pi_{1,0}N - (p_{R,R} + p_{O,R})R, \\ \frac{dR}{dt} &= p_{R,R}R, \\ \frac{dO}{dt} &= p_{O,R}R, \end{aligned} \quad (2)$$

The initial conditions are $X(0) = 0$, $R(0) = R_0 \geq 1$, $O(0) = 0$. The main output of this model is $dO/dt = p_{O,R}R$ and note that we are assuming that there is a fixed number of susceptible individuals, N , who may become members of the forum. We also need to take into account the fact that the success of a forum may depend on its popularity (for example, larger fora may be able to attract more new members). Thus, we make the assumption that $\pi_{1,0}$ has an



explicit dependence on the forum size as follows, $\pi_{1,0}(X) = p_N(X)X$, with $p_N(X) > 0$ for $X > 0$.

Next, we address the possibility that as fora become more popular, they are more likely to draw the attention of anti-extremist authorities and other agencies, and therefore may be attacked and shut down. We can model this by assuming that, when forum population (X) reaches a threshold θ , the forum is attacked and disabled, causing all its variables (X, R, O) to go to 0.

Forum administrators may try to balance two forces. On the one hand, they are aware that larger fora are more likely to be attacked, so they may try to control the number of users on the forum. On the other, they may want to maximize membership, making X as large as possible (i.e., just below the threshold θ).

For simplicity, suppose the transition rates to become a recruiter or terrorist are constant such that $p_{R,R} = p_R^0 > 0$ and $p_{O,R} = p_O^0 > 0$. Assume X^* is the target value. Let $X(t^*) = X^*$, where X^* is the target control condition for the number of members of the forum. The stationarity condition $dX(t^*)/dt = 0$ implies that

$$p_N(X^*) = \frac{(p_R^0 + p_O^0)R^*}{N X^*} = : p_N^*, \tag{3}$$

where $R^* = R(t^*)$. We extrapolate $p_N(X^*)$ linearly in X and R , that is,

$$p_N(X, R) = \begin{cases} \frac{p_R^0 + p_O^0}{N X^*} R + p_N^0(X^* - X), & \text{if } X < X^* + \frac{p_N^*}{p_N^0} \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where p_N^0 is a positive constant so that $p_N(X, R) \geq 0$ for all $X, R \geq 0$. Note that the stationarity condition causes the transition rate per time unit and forum participant p_N to depend on both X and R . Because control is exerted by individual recruiters we assume condition (4) will be applied with a time delay $\tau > 0$. This delay is one factor that makes it impossible for the recruiter to stay below the target X^* indefinitely (see Methods, section ‘‘Implications of Delayed Control’’).

We also need to take into account that the target set by the recruiters, X^* , cannot be stationary over time, because the lack of attacks on the forum raises the expectations of the target values. We model this by a linear factor, λ , that increases the target value, X^* over time. Figure 1 shows an example of the exponential growth rate of the forum and how it tends to decrease as the number of members increase. The equations for the controlled forum can be written as

$$\begin{aligned} \frac{dX}{dt} &= p_N(X(t-\tau), R(t-\tau), X^*)(N - X - R - O)X - (p_R^0 + p_O^0)R, \\ \frac{dR}{dt} &= p_R^0 R, \\ \frac{dO}{dt} &= p_O^0 R, \\ \frac{dX^*}{dt} &= \lambda, \end{aligned} \tag{5}$$

where $p_N(X, R, X^*)$ is given as in equation (4). Note that we have replaced N in the first equation by $(N - X - R - O)$ because the number of individuals susceptible of becoming forum members are the ones that remain outside the Forum.

It can be shown analytically that this model with control displays the double exponential behavior of the empirical data. There is fast exponential growth when forum administrators are not exerting control ($X(t) = Ae^{ct} - a/c$ when $t \approx \tau, t > \tau$), and slower exponential growth ($X(t) = Be^{pt} + q/p = Be^{pt} + X^*$ when $t \approx t^*$) when they are close to their control target, i.e. $c > p$ (see Methods, Section ‘‘Linear asymptotic behavior of the forum’’). Moreover, there is additional, subtle evidence in the data that such control mechanisms exist – notice in Figure 1 that Forum 5 breaks out of an equilibrium around

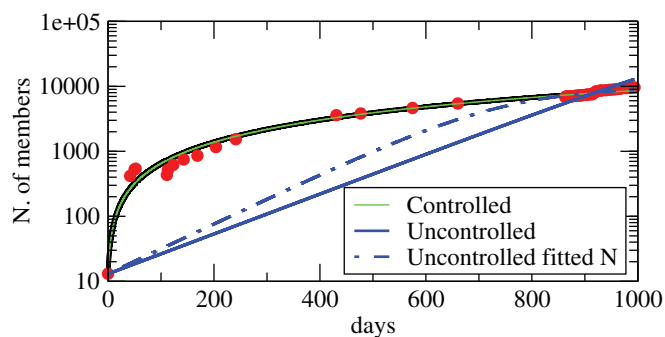


Figure 2 | Fit of the model to the Forum 8 by minimizing root mean square error (RMSE) for the model with control (solid black line) and the uncontrolled one (solid blue). Note that even when we add an additional parameter to the uncontrolled model (the number of susceptible individuals, N) the model without control cannot fit the data well (dashed line).

Day 1450 and becomes exponential again, suggesting that the control mechanism was discontinued at that point.

We also derived an alternative model without control in which the only limiting factor for growth of the forum was the size of the population N (see Methods, section ‘‘Specification of the model without control’’). In other words, it is possible that the slowdown in growth may be driven by the fact that there are fewer susceptible individuals in the population as forum membership increases. However, in Figure 2 we use data from Forum 8 to demonstrate that the model with control outperforms the alternative model, even when we let N be a free parameter of the model that we can fit. We also show in Table 3 the values fitted to other fora in our data set, all of which indicate that the controlled model fits considerably better than the uncontrolled model. The parameter p_N^0 can be easily estimated from the equations that describe the initial growth of the forum (see Methods, section ‘‘Linear asymptotic behavior of the forum,’’ CASE 1). The parameters p_R^0 and p_O^0 can also be uniquely estimated by the saturation process of the forum (details in Methods section ‘‘Linear asymptotic behavior of the forum,’’ CASE 2). The rest of the parameters are obtained by minimizing the RMSE using the PGAPack library⁵². Note that the mean-field model we propose is based on a deterministic (not stochastic) dynamical system, and therefore it has a unique solution for a particular set of parameters.

We now extend our model of a single forum to one that includes the interplay between several fora by linking them together in a model of competition. Let I be the total number of individuals, $N(t)$ the number of individuals susceptible to entering a forum (that is, those who have not yet entered a forum at time t), N_F the total number of fora. For simplicity let us assume that each individual can only belong to one forum. Then,

$$I = N(t) + \sum_{j=1}^{N_F} (X^j(t) + R^j(t) + O^j(t)), \tag{6}$$

which means that the total number of individuals in the system is conserved. When a particular forum j is eliminated at time τ , then $N(\tau_+) = N(\tau_-) + X^j(\tau_-)$ and $X^j(\tau_+) = 0$. Note that this may lead to a potential increase in the growth of other active fora. Forum j can reappear again (perhaps with a different name or identity) but with a lower equilibrium point, adjusted below the level that triggered the attack as $X^{j*}(\tau_+) = \mu X^{j*}(\tau_-)$, where μ can vary in the range $(0, 1)$. None of the fora know *a priori* what the threshold $\theta = X^j(\tau_-)$ value is, and this uncertainty creates a market for competition ruled by self-regulation. We model the reappearing time with a Poisson distribution with rate of 1 event in time T_R .



Table 3 | Parameter values of the model fitted to each individual forum by minimizing root mean square error (RMSE) for the controlled model. For comparison, the RMSE obtained after fitting the parameters of the uncontrolled model is also shown for each forum. Forum 5 is not included in the table as it breaks out of an equilibrium around Day 1450 and becomes exponential again, suggesting that the control mechanism was discontinued at that point

Forum Name	RMSE control	RMSE no-control	p_N^0	p_R^0	p_O^0	$X^*(t=0)$	λ
Forum 1	545	1239	$2.42 \cdot 10^{-8}$	0.002	0.0004	2275	19.58
Forum 2	97	106	$3.87 \cdot 10^{-8}$	0.002	0.000044	4153	0.09
Forum 3	85	171	$3.50 \cdot 10^{-8}$	0.0003	0.001	2937	12.40
Forum 4	29	271	$2.75 \cdot 10^{-8}$	0.002	0.0008	2220	0.66
Forum 6	49	122	$1.31 \cdot 10^{-8}$	0.005	0.0004	7430	10.38
Forum 7	73	141	$8.08 \cdot 10^{-8}$	0.005	0.0004	7423	10.44
Forum 8	436	1398	$2.66 \cdot 10^{-8}$	0.0002	0.0007	2330	10.83
Forum 9	48	122	$1.32 \cdot 10^{-8}$	0.005	0.0005	7432	11.57
Parameter Mean			$3.24 \cdot 10^{-8}$	0.0026	0.0006	4525	9.49
Standard Deviation			$2.00 \cdot 10^{-8}$	0.0020	0.0002	2480	6.37

We make use of the conservation identity in equation (6) to derive the following expressions for rates of growth in each forum $i = 1, \dots, N_F$:

$$\begin{aligned} \frac{dX^i}{dt} &= p_N(X^i(t-\tau), R^i(t-\tau), X^{i,*})X^i \times \\ &\left[I - \sum_{j=1}^{N_F} (X^j(t) + R^j(t) + O^j(t)) \right] - (p_R + p_O)R^i, \\ \frac{dR^i}{dt} &= p_R^0 R^i, \\ \frac{dO^i}{dt} &= p_O^0 R^i, \\ \frac{dX^{i,*}}{dt} &= \lambda^i. \end{aligned} \quad (7)$$

Initially each forum has zero members, $X^i(t=0) = 0$, and at least 1 recruiter, $R^i(t=0) = 1$, and produces no violent output $O^i(0) = 0$. Note that each forum fixes their own target values $X^{i,*}(t=0)$, with their own estimation of the threshold starting with a uniformly random distribution of thresholds at the beginning. Since fora are started at random times, we initialize $p_N(X^i(t-\tau), R^i(t-\tau), X^{i,*})$ with some probability, κ .

As an illustration we conduct a numerical simulation of 50 fora, $N_F = 50$, depicted in Figure 3. An external agent (for example, a counter-terrorism agency) allows fora to grow until they reach 14,000 members ($\theta = 14,000$) but no forum administrator knows this *a priori*. In this example, we assume that administrators apply their control rules with a one day delay, $\tau = 1$. When a forum is attacked and deactivated, its membership goes back to 0, and it reappears with an adjusted threshold that is $\mu = 0.99$ times the former value. We also assume that on average 20 fora can be activated in a year, that is $\kappa = 20/(N_F \times 365 \text{ days})$.

The results show that fora are able to control the size of their membership for long periods of time, which is not a trivial problem from a distributed control point of view. Moreover, differences in forum lifetimes are based on the heterogeneity in their different control target thresholds, $X^{i,*}$, and the variability in the parameter values p_N^0 , p_R^0 , and p_O^0 . This effect leads to different rates of growth for different fora, even though the fact that policymakers exert the same pressure on all of them.

To illustrate the implications of self-regulation, we show in Figure 4 the dependence of total output and the number of fora disrupted as a function of the policy maker threshold, θ . Notably,

total output increases nearly exponentially with the threshold, suggesting that even weak control mechanisms can have a big effect on reducing membership. However, the marginal effect of a stricter (lower θ) policy decreases dramatically as control is tightened. Note also that the number of fora disrupted decreases exponentially with θ , suggesting that efforts to tighten control will require an increasingly costly effort in detection and deactivation. Therefore, the amount of extremism reduced per unit cost of control will decline quickly with stricter policies both because of decreasing effectiveness of the attacks as well as increasing costs of detection/deactivation of fora.

Discussion

The model presented here reproduces the observed dynamics of violent extremist fora. In particular, the model shows (i) an initially steep and exponential increase in forum members, (ii) an intermediate interval in which the number of new members continues to grow exponentially but at a decreasing rate, and (iii) a final stabilization period, in which the number of participants smoothly converges (and, eventually, overshoots) a target number of members. Our analysis shows that the model is robust to varying assumptions about the recruitment and control process, accommodating any arbitrary number of status ranks in these fora, and allowing for time delay in the implementation of the access control mechanism.

Even though we and other researchers lack detailed evidence about when and how forum self-controls might be implemented, it is known that extremist fora are the target of law enforcement organizations. Our model suggests that a simple strategic reaction to these attacks can explain dynamic changes in forum membership. A plausible explanation suggested by field-researchers is that the forum leaders are able to 'broadcast' their message to the population in a top down manner rather than across a well defined network structure, which is known to display a bursty, slow speed of diffusion via word-of-mouth^{53,54}. Our model can also explain oddities in the empirical data, such as the observation that some fora stop growing for long periods of time, and, in some cases, they suddenly start growing after being stagnant (see Figure 1; Forum 5). This suggests that some fora occasionally eliminate or drastically alter their self-control mechanisms to allow for a new burst of growth.

The model also lets us test the effect of different strategies that might be employed by law enforcement agencies. The fact that membership increases exponentially with less strict policies means that stricter policies may be decreasingly effective at the margin. Moreover, the number of fora that must be attacked increases exponentially with stricter policies, suggesting that the costs of enforcing stricter policies may increase exponentially. Thus, we expect a sharp

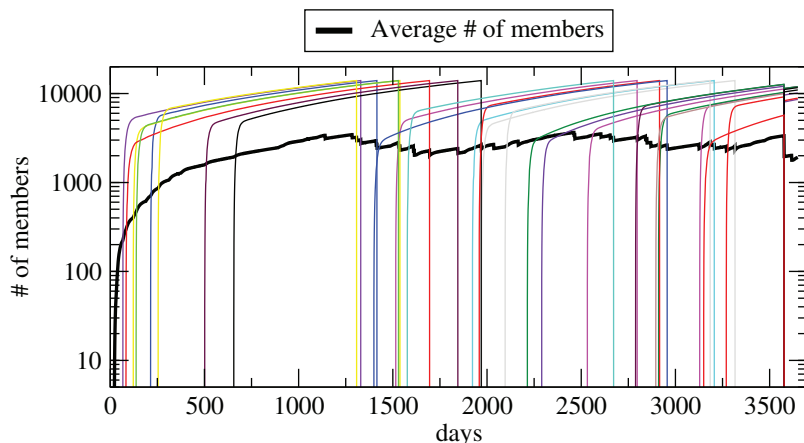


Figure 3 | Simulated number of members of 25 randomly chosen fora out of 50 during a period of 10 years: $I = 17,000,000$, $N_f = 50$, $p_N^0 = 3.24e - 8$, $p_R = 0.0026$, $p_O^0 = 0.0006$, $\theta = 14,000$, and $\tau = 1$ day, and $\lambda = 9.49$.

decline in the effectiveness of targeting additional fora of smaller size. As a result, we would recommend targeting policies aimed at occasionally disrupting a small number of the largest groups. Resources devoted to a stricter policy might be better used elsewhere. Our findings therefore back Berger’s “Strategy doesn’t have to be an all-or-nothing proposition.”⁵⁵

An important limitation of our model is that it is just a model – there may be other ways to explain the growth rates we observe. Our results yield better fit than a simple growth model without control, suggesting that the fora have recruitment procedures that control access in order to keep the number of members below a time-dependent threshold. But it is possible that other processes, particularly those based on recruitment within social networks (rather than random recruitment from the population) may also fit the data. We also note that the limited size of our dataset may hinder our efforts to validate the proposed model, but it is important to emphasize that these data are difficult to collect due ethical, experimental, and legal impediments to monitoring the activity of extremist fora. Finally, one might expect that a deliberate policy of limiting the rate of growth to be something that would itself be discussed on a forum, but during our observations we did not find any evidence of such discussions.

That said, if this model approximates the true underlying process, it suggests that even minimal efforts to target extremist fora may have a large effect on reducing their membership and therefore their capacity to recruit for real world violent acts. It also suggests that much of this effect is due to the self-control induced in the fora rather than the direct effect of temporarily dispersing the members of a given forum.

While our original aim was to better understand extremist groups, we note that our model might also be of use to those seeking to

understand how legitimate fora react under stress. For example, there have been numerous recent examples of repressive regimes that seek to thwart the use of social media to mobilize legitimate opposition protest^{56,57} and in future work it would be interesting to see whether our model can explain changes in the membership and self-control mechanisms adopted by fora that are used in such protests.

Methods

Forum names and URLs are available from the authors on request.

Implications of delayed control. The evolution equation for X is

$$\begin{aligned} \frac{dX}{dt} &= p_N(X(t-\tau), R(t-\tau))N X(t) - (p_R^0 + p_O^0)R(t) \\ &= \frac{p_R^0 + p_O^0}{X^*} R(t-\tau)X(t) + p_N^0 N [X^* - X(t-\tau)]X(t) - (p_R^0 + p_O^0)R(t) \end{aligned} \tag{8}$$

for $X \leq X^* + p_N^*/p_N^0$, and

$$\frac{dX}{dt} = -(p_R^0 + p_O^0)R \tag{9}$$

for $X > X^* + p_N^*/p_N^0$. The resulting model yields delayed control behavior that is like steering a boat: one tends to overshoot the target values because the action is implemented with a delay.

Since the smaller τ the better the control, we will consider the condition expressed in equation (8). The target value X^* would be reached at time t^* by the uncontrolled system. At this point,

$$\begin{aligned} R(t^* - \tau) &= R^* - \tau \frac{dR}{dt}(t^*) + O(\tau^2) \\ &= (1 - \tau p_R^0)R^* + O(\tau^2) \end{aligned} \tag{10}$$

by (5), and

$$\begin{aligned} X^* - X(t^* - \tau) &= \tau \frac{dX}{dt}(t^*) + O(\tau^2) \\ &= \tau p_N^0 N X^* - \tau (p_R^0 + p_O^0)R^* + O(\tau^2) \\ &= \tau (p_R^0 + p_O^0)R^* - \tau (p_R^0 + p_O^0)R^* + O(\tau^2) \\ &= O(\tau^2) \end{aligned} \tag{11}$$

by equation (2) and equation (3). It follows from equation (8) that up to $O(\tau^2)$,

$$\frac{dX}{dt}(t^*) = (p_R^0 + p_O^0)(1 - \tau p_R^0)R^* - (p_R^0 + p_O^0)R^* + O(\tau^2) = -\tau (p_R^0 + p_O^0)p_R^0 R^* + O(\tau^2).$$

The fact that $dX(t^*)/dt < 0$ for $\tau > 0$ sufficiently small, shows that $X(t)$ exceeds X^* before the time $t = t^*$ (since $X(t^* - \tau) \simeq X^* - \tau \frac{dX}{dt}(t^*) > X^*$).

Linear asymptotic behavior of the forum. Let us linearize at the early stages of the formation of the forums and at later stages when forum membership reaches a high number. We know from Eq. (5) that

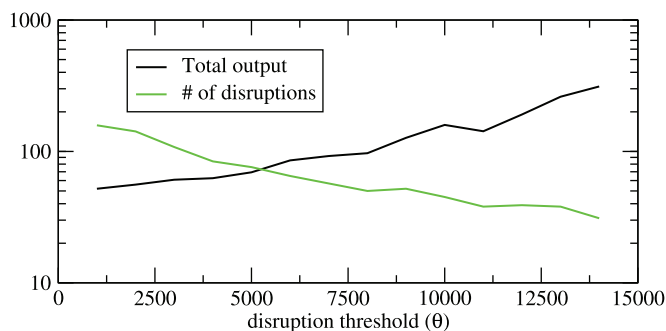


Figure 4 | Total output and number of forum disruptions versus the policymaker threshold θ (lower values indicate stricter policies) using the same set of parameters as in Figure 3.



$$R(t) = R_0 e^{p_R^0 t}$$

Replacing this in Eq. (8), we get

$$\frac{dX}{dt}(t) = \frac{p_R^0 + p_O^0}{X^*} R_0 e^{p_R^0(t-\tau)} X(t) + p_N^0 N [X^* - X(t-\tau)] X(t) - (p_R^0 + p_O^0) R_0 e^{p_R^0 t}. \quad (12)$$

This equation is nonlinear due to the second term on the right hand side. With this equation we can now study two limiting cases.

CASE 1. If $t \approx \tau$, $t > \tau$, then we may replace $X^* - X(t - \tau)$ by X^* in (12) since $X(t - \tau) \ll X^*$. Eq. (12) becomes linear:

$$\frac{dX}{dt}(t) - P(t)X(t) = Q(t), \quad (13)$$

where

$$P(t) = \frac{p_R^0 + p_O^0}{X^*} R_0 e^{p_R^0(t-\tau)} + p_N^0 N X^* \approx p_N^0 N X^* = : c > 0, \quad (14)$$

because $e^{p_R^0(t-\tau)} \approx 1$ and $R_0/X^* \ll 1$, and

$$Q(t) = -(p_R^0 + p_O^0) R_0 e^{p_R^0 t} \approx -(p_R^0 + p_O^0) R_0 = : -a (a > 0).$$

Hence

$$X(t) = A e^{ct} - \frac{a}{c}.$$

CASE 2. If $t \approx t^*$, then $X^* - X(t - \tau) \approx 0$ according to Eq. (11). Eq. (12) becomes also linear, see Eq. (13), but now

$$P(t) = \frac{p_R^0 + p_O^0}{X^*} R_0 e^{p_R^0(t^*-\tau)} \approx \frac{p_R^0 + p_O^0}{X^*} R_0 e^{p_R^0 t^*} = : p > 0, \quad (15)$$

$$Q(t) = -(p_R^0 + p_O^0) R_0 e^{p_R^0 t} \approx -(p_R^0 + p_O^0) R_0 e^{p_R^0 t^*} = : -q = -pX^*.$$

Analogously to Case 1, the solution is

$$X(t) = B e^{pt} + \frac{q}{p} = B e^{pt} + X^*.$$

This shows that the population dynamics model is consistent with the double exponentials observed in the data. Moreover, it follows from Eq. 14 that $X(t)$ starts growing at an exponential rate given by the parameter

$$c = p_N^0 N X^*.$$

By the time $X(t) \lesssim X^*$, the exponential growth rate implied by Eq. 15 is

$$p = (p_R^0 + p_O^0) \frac{R^*}{X^*} \ll p_R^0 + p_O^0.$$

Since $N, X^* \gg 1$, and on general grounds $p_N^0 > p_R^0, p_O^0$, we conclude that $c > p$, i.e., the growth of forum members is much greater at the beginning than at the end of the time span $(0, t^*)$.

Specification of the model without control. The model without control can be simply written as

$$\frac{dX}{dt} = p_N(N - X - R - O)X - (p_R^0 + p_O^0)R,$$

$$\frac{dR}{dt} = p_R^0 R,$$

$$\frac{dO}{dt} = p_O^0 R.$$

These equations show that the only mechanism the uncontrolled system has for saturation is the reduction of susceptible individuals $(N - X - R - O)$ over time. However Table 3 and Figure 2 indicate that even when N is small and an adjustable parameter, the controlled model fits the empirical data considerably better than the uncontrolled model.

1. Newman, M. The structure and function of complex networks. *SIAM review* **45**, 167–256 (2003).
2. Watts, D. The “new” science of networks. *Annual Review of Sociology*, 243–270 (2004).
3. Kleinberg, J. The convergence of social and technological networks. *Communications of the ACM* **51**, 66–72 (2008).

4. Lazer, D. *et al.* Social science: computational social science. *Science* **323**, 721 (2009).
5. Barabási, A. The network takeover. *Nature Physics* **8**, 14 (2011).
6. Bohorquez, J., Gourley, S., Dixon, A., Spagat, M. & Johnson, N. Common ecology quantifies human insurgency. *Nature* **462**, 911–914 (2009).
7. Clauset, A. & Gleditsch, K. The developmental dynamics of terrorist organizations. *PLoS ONE* **7**, e48633 (2009).
8. Johnson, N. *et al.* Pattern in Escalations in Insurgent and Terrorist Activity. *Science* **333**, 81 (2011).
9. Bohannon, J. Counting the Dead in Afghanistan. *Science* **331**, 1256 (2011).
10. Hicks, M., Lee, U., Sundberg, R. & Spagat, M. Global Comparison of Warring Groups in 2002–2007: Fatalities from Targeting Civilians vs. Fighting Battles. *PLoS one* **6**, e23976 (2011).
11. Lim, M., Metzler, R. & Bar-Yam, Y. Global pattern formation and ethnic/cultural violence. *Science* **317**, 1540 (2007).
12. Nelson, G. Violence: the role of society and state. *Nature* **480**, 181–181 (2011).
13. Hsiang, S., Meng, K. & Cane, M. Civil conflicts are associated with the global climate. *Nature* **476**, 438–441 (2011).
14. Rutherford, A., Harmon, D., Gard-Murray, A., Gros, A. & Bar-Yam, Y. Good Fences: The Importance of Setting Boundaries for Peaceful Coexistence. Available at SSRN 1942515 (2011).
15. Atran, S. *et al.* Sacred barriers to conflict resolution. *Science* **317**, 1039–1040 (2007).
16. Ginges, J., Atran, S., Medin, D. & Shikaki, K. Sacred bounds on rational resolution of violent political conflict. *Proceedings of the National Academy of Sciences* **104**, 7357 (2007).
17. Apic, G., Betts, M. & Russell, R. Content Disputes in Wikipedia Reflect Geopolitical Instability. *PLoS one* **6**, e20902 (2011).
18. Morselli, C., Giguere, C. & Petit, K. The efficiency/security trade-off in criminal networks. *Social Networks* **29**, 143–153 (2007).
19. Lindelauf, R., Borm, P. & Hamers, H. The influence of secrecy on the communication structure of covert networks. *Social Networks* **31**, 126–137 (2009).
20. Sagarin, R. *et al.* Decentralize, adapt and cooperate. *Nature* **465**, 292–293 (2010).
21. Weimann, G. *Terror on the Internet: The new arena, the new challenges* (United States Institute of Peace, 2006).
22. Lia, B. Al-Qaeda online: understanding jihadist internet infrastructure. *Janes Intelligence Review* **18**, 14 (2006).
23. Efthymiopoulos, M. & Demergis, J. In *Proceedings of 6th European Conference on Information Warfare and Security* (2007), 341.
24. Flade, F. Death of a online-jihadi: from cyberspace to battlefield. *News Of Terrorism, Jihadism & International Politics* (December 2011).
25. Rogan, H. Jihadism Online: A study of how al-Qaida and radical Islamist groups use the Internet for terrorist purposes. *FFI/Report, Norwegian Defence Research Establishment* **915** (2006).
26. Brachman, J. & Levine, A. The World of Holy Warcraft. *Foreign Policy* (2011).
27. Allievi, S. Islam in the public space: social networks, media and neo-communities. *Muslim Networks and Transnational Communities in and across Europe*, 1–27 (2003).
28. Kohlmann, E. The Real Online Terrorist Threat. *Foreign Affairs* **85**, 115 (2006).
29. Ali Musawi, M. Cheering for Osama: How jihadists use internet discussion forums. *New Quilliam Report* (September 2010).
30. Brachman, J. High-tech terror: Al-Qaeda’s use of new technology. *Fletcher Foreign World Affairs* **30**, 149 (2006).
31. Middle East Media Research Institute (<http://www.memri.org>). Accessed 2012/2/1.
32. NEFA Foundation (<http://www.nefafoundation.org>). Accessed 2012/2/1.
33. SITE Intelligence Group (<http://news.siteintelgroup.com>). Accessed 2012/2/1.
34. Internet Haganah (<http://internet-haganah.com>). Accessed 2012/2/1.
35. ICT’s Jihadi Websites Monitoring Group (<http://www.ict.org.il>). Accessed 2012/2/1.
36. Jihad Watch (<http://www.jihadwatch.org/>). Accessed 2012/2/1.
37. Jordan, J., Torres, M. & Horsburgh, N. The Intelligence Services’ Struggle Against al-Qaeda Propaganda. *International Journal of Intelligence and CounterIntelligence* **18**, 31–49 (2004).
38. Conway, M. Terrorism and the Internet: New Media—New Threat? *Parliamentary Affairs* **59**, 283 (2006).
39. Hegghammer, T. Spy forum (<http://www.jihadica.com/spy-forums/>) Accessed 2012/2/1. (2010).
40. Nakashima, E. Dismantling of saudi-cia web site illustrates need for clearer cyberwar policies. Available at: <http://www.washingtonpost.com/wp-dyn/content/article/2010/03/18/AR2010031805464.html>. Accessed 2012/2/1. *Washington Post* (2010).
41. Musick, J. Counterstrike: the untold story of America’s secret campaign against Al Qaeda. *Journal of Policing, Intelligence and Counter Terrorism* **7**, 180–182 (2012).
42. Torres, M. The Vulnerabilities of Online Terrorism. *Studies in Conflict & Terrorism* **35**, 263–277 (2012).
43. Abend, L. Who is waging cyberwar against the jihadi networks? Available at: <http://www.time.com/time/world/article/0,8599,2111293,00.html>. Accessed 2012/12/1. *Time* (2012).
44. Katz, R. & Stephenson, W. Terrorist on the Internet. Available at: <http://theconnection.wbur.org/2005/06/08/terrorist-on-the-internet>, Accessed 2012/22/2. *The Connection* (2005).



45. Juzgado Central de Instrucción #3; Diligencias previas 126/201; Available at: http://www.elpais.com/elpaismedia/ultimahora/media/201108/20/espana/20110820elpepunac_1_Pes_PDF.pdf. Accessed 2012/2/1. (2011).
46. Anderson, R., May, R. & Anderson, B. *Infectious diseases of humans: dynamics and control* (Wiley Online Library, 1992).
47. Diekmann, O. & Heesterbeek, J. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation* (Wiley, 2000).
48. Rogan, H. Al-Qaedas online media strategies: From Abu Reuter to Irhabi 007. *FFI/Report, Norwegian Defence Research Establishment* **2729** (2007).
49. Homan, R. The ethics of covert methods. *British Journal of Sociology* **31**, 46–59 (1980).
50. Herrera, C. Two arguments for covert methods in social research. *The British journal of sociology* **50**, 331–343 (1999).
51. Satter, R. Sarkozy: Jail those who browse terror websites. Available at: <http://news.yahoo.com/sarkozy-jail-those-browse-terror-websites-182328946.html>. Accessed 2012/12/1. *Associated Press* (2012).
52. Levine, D. Users guide to the PGAPack parallel genetic algorithm library. *Argonne National Laboratory* **9700** (1996).
53. Iribarren, J. & Moro, E. Impact of human activity patterns on the dynamics of information diffusion. *Physical review letters* **103**, 38702 (2009).
54. Karsai, M. *et al.* Small but slow world: how network topology and burstiness slow down spreading. *Physical Review E* **83**, 025102 (2011).
55. Berger, J. #Unfollow. Available at <http://www.foreignpolicy.com/articles/2013/02/20/unfollow>. Accessed 2012/22/2. *Foreign Policy* (2013).
56. Wright, S. Government-run Online Discussion Fora: Moderation, Censorship and the Shadow of Control. *The British Journal of Politics & International Relations* **8**, 550–568 (2006).
57. King, G., Pan, J. & Roberts, M. How Censorship in China Allows Government Criticism but Silences Collective Expression. Available at: <http://gking.harvard.edu/sites/scholar.iq.harvard.edu/files/gking/files/censored.pdf>. Accessed 2012/22/2. *APSA 2012 Annual Meeting Paper* (2012).

gking.harvard.edu/sites/scholar.iq.harvard.edu/files/gking/files/censored.pdf. Accessed 2012/22/2. *APSA 2012 Annual Meeting Paper* (2012).

Acknowledgments

James H. Fowler acknowledges support from the James S. McDonnell Foundation grant #220020177/SC. Manuel Cebrian acknowledges support from the National Science Foundation under grant 0905645, from DARPA/Lockheed Martin Guard Dog Program under PO 4100149822, and the Army Research Office under Grant W911NF-11-1-0363. Manuel R. Torres was supported by The Spanish Ministry of Science and Technology for his time in residence at the Weatherhead Center for International Affairs, Harvard University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Manuel Cebrian, Manuel R. Torres, Ramon Huerta and James H. Fowler collected the data, designed and performed the research as well as wrote the paper.

Additional information

Competing financial interests: The authors declare no competing financial interests.

License: This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

How to cite this article: Cebrian, M., Torres, M.R., Huerta, R. & Fowler, J.H. Violent extremist group ecologies under stress. *Sci. Rep.* **3**, 1544; DOI:10.1038/srep01544 (2013).