

The Hate Speech and Discussion Topics of White Supremacists During the Pandemic

Anh V. Vu

10 November 2020

Executive summary

CCC COVID Briefing Papers are an ongoing series of short-form, open access reports aimed at academics, policymakers, and practitioners, which aim to provide an accessible summary of our ongoing research into the effects which the coronavirus pandemic (and government responses) are having on cybercrime.

In this report, we explore one of the largest extremist, white supremacist, online forums: Stormfront. We analyse discussion topics and hate speech related to the pandemic in a subset of approximately 15 000 posts from our ExtremeBB dataset, which were scraped from a board on this forum dedicated to discussions about the pandemic. We find discussion topics include *mask-wearing*, *suspicious*, *ethnicity*, *politics*, *virus*, and *vaccine*. Additionally, by running an NLP classifier over those posts, we find 7.4% of them contain hate speech and offensive language; these discussions are, specifically, deeply antisemitic, including baseless conspiracy theories as to the virus's origin and alleged intended political consequences.

Stormfront and white supremacy

Stormfront is one of the largest and longest-running (since 1996) online platforms of the extreme-right movement. It is an online gathering place for white supremacists, racists, antisemites, and right-wing extremists [1]. Some of its users have been involved in serious illegal harm, and the site is associated with a number of terrorist atrocities, including the 2012 Wisconsin Sikh temple shooting, the Pittsburgh police officer shooting in 2009, and notoriously, the mass killing of 77 people in Norway in 2011. In response to the pandemic, Stormfront opened a dedicated board for Coronavirus discussion in January 2020. We examine posts from this board to shed light on extremists' thoughts during the crisis.

Popular discussion topics

We use Latent Dirichlet Allocation to discover hidden topics for the collected posts, then manually label them by looking at each topic's keyword distribution. The main theme of a post is then assigned by its dominating topic – the one that has highest probability among its topic distribution. To select the optimal model, we trained several models, spanning 10 to 100 topics, then decided the number of topics based on the coherence score. Our experiment shows a 10-topic model is optimal. We find the most popular topic is *mask-wearing* accounting for 13.5% of posts. The next most popular topics are *suspicious* (12.1%), *multimedia* (11.4%), *ethnicity* (11.1%), and *politics* (10.0%). We also see discussion regarding *virus* and *vaccine* with 8.2% and 6.6% of posts, respectively. We then evaluate the popularity evolution of those topics, showing the ranking of top popular topics *all* increased after the lockdown in March 2020

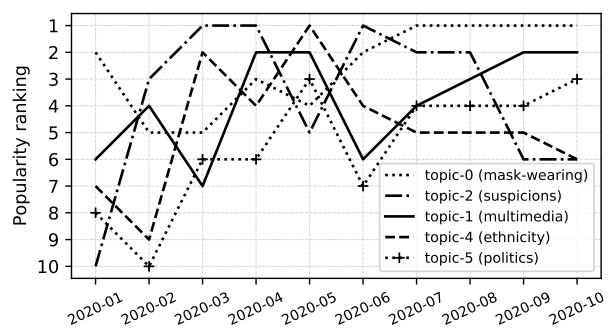


Figure 1: The popularity shift of topics

(Figure 1), particularly *politics* jumped from rank #10 just before the lockdown in February 2020 to rank #3 at the end of the period. Another big leap is of *ethnicity*, which increased from rank #9 to #2 in response to the lockdown, and ranked #1 in May 2020 before it gradually decreased.

We also examine how these popular topics are distributed between groups of members. Around 300 members were involved in each of those topics, with each member, on average, contributing around 6 posts. The intersection of members between each pair of those topics is around 200, and is nearly 130 between *all* topics. This suggests a fairly diffuse discussion structure; most of the members tend to be involved in a wide range of discussion topics, rather than distributed among smaller groups which focus on a subset of topics.

Hate speech in light of the pandemic

Hateful comments on online forums are often more aggressive due to poster anonymity, and such forums facilitate the spreading of extreme material on the Internet despite blocking by sites such as Facebook and Twitter. We use a multi-class classifier proposed by Davidson et al. [2] to discover forum posts containing hate speech and offensive language, finding that the number of such posts and the number of posters increased sharply in the first few months after the WHO declared the global pandemic, particularly in March and April 2020 when we observe a high, but fairly short-lived, peak in volume (Figure 2). Afterwards, posting activities subsided, showing the lockdown fueled hate discussions, but just for a short period, presumably while the pandemic was front and centre in everyone's attention.

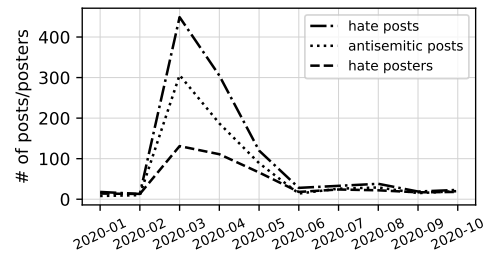


Figure 2: Volume of posts and posters

The most common hate contents are antisemitic, which accounts for the largest proportion of hate posts and also for the peaks in March 2020 (see Figure 2). Discussions of the pandemic revolve around antisemitic conspiracy theorising – insisting that a group of powerful conspirators created and then spread the deadly virus. When the UK Prime Minister tested positive, his (potential) death and the consequences were also discussed, with forum members suspecting the virus was a conspiracy to kill him and let a Jewish take power. Some also believe in a conspiracy that the virus is a human-crafted and the disease is intentionally targeted.

Conclusions

The lockdown earlier in 2020 intensified hate speech on a board dedicated to Coronavirus on Stormfront. Although the prevalence of antisemitism on an extreme-right discussion board is hardly surprising, the short-lived nature of this intensification of hate speech volumes is worthy of comment, suggesting, as for our cybercrime collections, that this increase may be as much due to changes in routine activities of frequent posters on these forums (i.e. increased free time under lockdown) as to reaction to the pandemic (which has only intensified in the US across the summer). We will be monitoring how hate speech and discussion topics evolve during the second phase of lockdowns now occurring in late 2020. We predict another peak, but time will tell whether or not it is more pronounced.

[1] L. Bowman-Grieve (2009). Exploring “Stormfront”: A Virtual Community of the Radical Right. *Studies in Conflict & Terrorism*, 32(11), 989–1007.

[2] T. Davidson, D. Warmley, M. Macy, I. Weber (2017). Automated Hate Speech Detection and the Problem of Offensive Language. *The International AAAI Conference on Web and Social Media (ICWSM)*.

At the Cambridge Cybercrime Centre we make our research data available to other academics, sometimes before we have looked at it ourselves! Researchers can be provided access to our ‘CrimeBB’ dataset of (26 and counting) underground cybercrime forums, our extensive collections of chat channel data, and our new collections of forums relating to online right-wing extremism and radicalisation. We can also share email spam and sensor data related to DDoS and IoT malware. All these collections are regularly updated and can be rapidly provided under licence – for full details see: <https://cambridgecybercrime.uk>

The full set of CCC COVID Briefing Papers can be found at: <https://cambridgecybercrime.uk/COVID>

This work is licensed under CC BY 4.0. To view a copy of this licence visit: <https://creativecommons.org/licenses/by/4.0>