**ABSTRACT:**

YouTube is the largest and most popular video hosting platform in the world with over 30 million visitors and almost 5 billion videos watched per day. It has been hypothesized by researchers and media that a radicalization pathway exists on this platform that draws normal users towards extreme political content. In this paper, we sample over 800 channels and over 2400 videos of both independent YouTube creators alongside organizations that post content. These channels are then grouped by their political ideology and specific category tags, and videos and their recommended videos are then sampled from these groupings to analyze. The sampled videos were then used to develop a Markov model that paints a picture of how these groupings interact with one another. Our findings contradict the popular belief of a radicalization pipeline, and suggests that YouTube has a deradicalizing nature, favoring center leaning and a few right-wing groups. We also analyzed views and channels from each ideology and tag to gain further insight on user watch preference based on video popularity.

* Maybe include more information about analyzing videos

**KEYWORDS:**

Radicalization; YouTube; Recommendation Algorithm; Algorithmic Extremism

**INTRODUCTION:**

* Introduces the topic of study and talks about information related to it (I expect a mini paragraph for each topic)
  + Where has it popped up?
  + How does the algorithm work?
  + How does YouTube handle radical content?
  + A look into YouTube’s Community Guidelines
  + YouTube Drama
  + How does other media handle it?
  + Effects of the radicalization algorithm and their recommendations
  + Prior History
  + Clickbait titles

Introduction and COVID-19

* With the COVID-19 pandemic closing many popular gathering places and forcing educational facilities and workplaces to transition towards a Work-From-Home model, internet usage has soared with steaming services increasing by 12% as well as overall internet usage increasing by 70% (CITE 1). This transition has increased the daily use of social media as well as has provided an opportunity for conspiracy theorists and the alt-right to thrive with COVID-19 related conspiracy theories. These conspiracies range in severity and create ties between COVID-19 with the introduction of new technology and politics, they claim that COVID-19 is a hoax, that it is connected to the introduction of 5G technology, as well as claiming that Americans aligned with the Democratic Party are overexaggerating its effects to hurt the Republican Party’s 2020 Presidential Election bid. (CITE 2)

Connecting to YouTube

* YouTube-type content has shown a 5% growth with at least 70% of internet users streaming videos daily. This growth amounts to 5 billion videos watch daily with an average viewing session of 40 minutes (CITE 3,4). The platform allows for a wide range of content from both individual creators and organizations and has seen a considerable growth in political reports, opinions, and conspiracy theories as its userbase grows over time.

YouTube Algorithm and Rabbit Hole

* The YouTube homepage contains an endless list of recommended videos to watch next, these videos include previously watched videos, but primarily suggests new videos that it thinks the user would like. Videos are generated dynamically, and new suggested videos are loaded as users reach the end of the screen. The homepage also includes general recommended topics (eg. Cooking, Comedy, The Sims 4) that the algorithm thinks the user would enjoy. Each recommended topic includes a finite list of previous watched and new videos that it fits under that topic. This recommendation system is highly personalized and more details about the algorithm is discussed in the following section. On the individual video level, when a user is watching a video, YouTube auto generates recommended videos to that topic as well as generates user specific recommendations titled ‘Recommended for you’ that are not related to the video. These recommendations as well includes previously watched videos as well as new unwatched videos. New recommendations generate as users scroll down through the webpage and stops after reaching a threshold that differs per video.
* YouTube’s video recommendation algorithm is based on numerous factors including user watch time, previous videos watched, metadata, and a multitude of other factors, but this algorithm affects user video consumption as over 70% of time spent on YouTube is watching videos recommended by the algorithm (CITE 5). The inner workings of the algorithm are unknown and it has been the site of heavy criticism with claims that the ‘autoplay’ feature where a new video plays once the video has finished has a tendency to recommend a more extreme version of videos currently played. In a report by the Atlantic, videos about vegetarianism led to veganism, and jogging led to running ultramarathons, suggesting that the algorithm has a bias towards inflammatory content (CITE 6). This claim then paints a picture of radicalization on the YouTube platform, where users starting out in partisan content, would then be recommended content from extreme political community, thus going down into a ‘rabbit hole’ where the content progressively becomes more extreme due to the algorithm’s recommendation (CITE 7)

YouTube Community Guidelines + Demonetization

* YouTube does have a strict set of rules about what could be posted on their platform, named the Community Guidelines. These rules aim at preventing violent or dangerous content including harassment, hate speech, and violent and/or graphical content, but videos falling under these categories are difficult for an automatic algorithm to detect (CITE 8). The YouTube Partner program also allows users to be monetized with ads provided that their content is deemed advertiser friendly. This program blocks ads on videos that include inappropriate language, hateful content, controversial issues and sensitive events to prevent the user from gaining revenue on their platform. (CITE 9)

Tie in to Paper and related work

* In this paper, we address the claims that there exists a radicalization pipeline on YouTube through a quantitative model of consumer behavior, specifically looking at YouTube videos and their recommendations. Our data gathering procedure and code is transparent and data sets will be available to the public for examination. More specifically, we ask:

RQ1: Does the YouTube recommendation algorithm steer users towards extreme content?

RQ2: How frequent does YouTube provide recommendations outside of the original channel’s ideology or category tag?

RQ3: How accessible are other ideologies/category starting from a specific ideology/category?

Using a dataset labelling political ideologies for over 800 channels, we sample videos uniformly at random on the 500 most recent videos on each channel, then we gather data about each video and collect their recommended videos. We analyze this collected data to assess the video views and popularity for each ideology/category tags, as well we aggregate the data to create a Markov model to model video consumption with the tags.

1. Beech, M. (2020, March 26). COVID-19 Pushes Up Internet Use 70% And Streaming More Than 12%, First Figures Reveal. Retrieved from <https://www.forbes.com/sites/markbeech/2020/03/25/covid-19-pushes-up-internet-use-70-streaming-more-than-12-first-figures-reveal/#b84d3153104e>
2. Adam M. Enders, J. E. (2020, May 05). Perspective | Conspiracy theories run rampant when people feel helpless. Like now. Retrieved July 13, 2020, from <https://www.washingtonpost.com/outlook/2020/05/05/coronavirus-conspiracy-theories-pandemic/>
3. Internet usage increased during March Covid-19 lockdown. (2020, June 19). Retrieved July 13, 2020, from <https://www.rte.ie/news/business/2020/0619/1148430-cso-internet-usage/>
4. Rutnik, M. (2019, August 09). YouTube in numbers: Monthly views, most popular video, and more fun stats! Retrieved from <https://www.androidauthority.com/youtube-stats-1016070/>
5. Cooper, P. (2019, November 01). How Does the YouTube Algorithm Work? A Guide to Getting More Views. Retrieved July 13, 2020, from <https://blog.hootsuite.com/how-the-youtube-algorithm-works/>
6. Friedersdorf, C. (2018, March 12). YouTube Extremism and the Long Tail. Retrieved July 13, 2020, from <https://www.theatlantic.com/politics/archive/2018/03/youtube-extremism-and-the-long-tail/555350/>
7. Roose, K. (2019, June 08). The Making of a YouTube Radical. Retrieved July 13, 2020, from <https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html>
8. Royer, A., & By. (n.d.). Why Youtube's Decision to Remove Far-Right Content is Not Enough. Retrieved July 13, 2020, from <https://www.themantle.com/arts-and-culture/why-youtubes-decision-remove-far-right-content-not-enough>
9. Advertiser-friendly content guidelines - YouTube Help. (n.d.). Retrieved July 13, 2020, from https://support.google.com/youtube/answer/6162278?hl=en

**RELATED WORK AND PRIOR STUDIES:**

* Mini paragraphs about the 3 drobox papers + other papers (aim for around 3-6?)
* Ledwich and Zaitsev
  + Youtube provides a deradicalization behavior favoring centralist ideals
* Rebeiro et Al.
  + Analyzed youtube comments from from 2006-2018 to see the frequency of people’s comments in different communities
* Zannettou
  + Analyzed comments from 4chan’s /pol/ and Gab to tie in frequency of words and historic events
* Read one more paper and summarize

**DATA COLLECTION AND MODELLING:**

* Data Collection Details
  + Used Mark Ledwitch’s TAGS, he used part from Rebeiro as well
  + Describe what each tag means
  + What did we collect from each channel?
  + What did we collect from each video?
  + How did we do the sampling?
  + Missing Videos and null values
  + “Others” TAG/Ideology discussion
  + Using a VPN? (Process at later date)
  + How did we collect each video? (how did we choose a timeframe)
  + What videos did we exclude? (10k views+)
* Modelling Details
  + Markov
    - What are impressions?
    - How did we weigh our probabilities?
    - What about others?

Introduction to Classification

* The Classification of YouTube Channels is no easy task, as YouTube content creators do not explicitly fall under an individual political ideology. There is no comprehensive list available tagging all YouTube content creators to a specific category, with multiple existing lists clashing with one another. We analyzed previous academic research studying radicalization on YouTube and selected the data set created by Ledwich and Zaitsev (2019). This data set aggregated data together from Ad Fontes Media and Media Bias Factcheck which categorizes popular political media and assesses their bias, combining these two sources covers almost 80% of all YouTube views and creates a relatively reliable categorization process. Ledwich and Zaitsev (2019) as well manually analyzed a list of channels along with a volunteer labeler where they assigned category tags for each individual video. When two or more labelers defined a channel by the same tag, that category tag is then assigned to the video, with a maximum of 4 category tags for each video.
* This Data was then aggregated together to roughly generalize thirteen political ideologies as shown by (TABLE #). We then sampled videos uniformly at random from both the category tags and ideology labels to gather their recommended video data to fit into our Markov model.

Data Collection + simulation process

* With this tagged dataset, we now describe our sampling technique. We begin by filtering the data set to each individual category tag to get a clearer view of the YouTube channels clustered together. We then built a web scraper that uses this channel data to extract 200 of the most recent videos with over 10,000 views from each individual channel. This provides us with a glimpse of the recent activity by the channel that has a substantial amount of views. This 10,000 views threshold was chosen to remove recently created content, which might provide a bias for each tag.
* We then chose 100 videos from the list of collected videos to gather data on, which provides us with a comprehensive view of each community. In this simulation process, each video is visited iteratively, and data on the video title, view count, as well as its first 10 recommended videos titles and channels are gathered. Only the first 10 recommended videos are selected as those videos are prominently featured on the sidebar as the user watches the current video.
* Since an individual channel could have 0 to an upwards of 4 tags, in this process we sampled each channel with more than 1 tag and for every tag that the channel is apart of. In this case, multiple channels are belonging to more than one community. In the case with the Ideology labels, since only one label is applied per channel, no extra steps are necessary.
* This data collection and simulation process is performed on the tags and ideologies that Ledwich and Zaitsev (2019) assigned. For the tagged data we sampled 800 channels and simulated 2476 videos for our analysis and for the ideology data we sampled 441 Channels with 1394 videos.

**FINDINGS AND DISCUSSIONS:**

Graphs and general findings:

* With the collected data we first analyze the viewership within each community to assess popularity for each tag. Figure 2 provides a scatterplot and a boxplot of views for each individual video within their respective tags providing a range from 0 views to 1 Million views, specifically with videos over 10,000 views in the data set. This boxplot suggests that mainstream media primarily dominates the political sphere of YouTube views, with Late Night Talk Shows, TV, and Educational channels hosting majority of the views while IDW, Alt Light, as well as Anti SJW channels maintain a middle-upper level standing in terms of viewership within the political sphere. (VIEWER COUNT)
* We then analyze how each community interacts within another using an impressions system. This system analyzed the number of times a channel is recommended within a particular YouTube video (CITE 1). In this study we are interested in the average impressions a YouTube video sends to different communities, this calculation is done by multiplying the current views the channel has by the number of times a YouTube channel is recommended divided by 10, this calculation is performed on each video analyzed, and then it is aggregated together for each tag to determine the average impressions tag A sends to other tags. (SANKEY DIAGRAMS)
* Figure 3 visualizes this data through a Sankey network diagram displaying the average flow of impressions between communities. Since channels are able to possess multiple tags, one tag is selected at uniform random to represent that channel. From this, we have high interactivity within mainstream communities in Educational, Late Night Talk Show, and TV channels sending large amounts of impressions to each other.
* Here we can analyze which channels are disadvantaged by the recommendation system by comparing the difference between impressions sent and impressions received as seen in figure 4. A positive advantage would indicate that that on average a tag has been recommended more times than it has sent out impressions.

Markov Model:

**LIMITATIONS /FUTURE DIRECTIONS AND CONCLUSIONS:**

1. Zarzycki, N. (2019, October 30). Reach vs. Impressions: What's More Important to Track? Retrieved July 15, 2020, from https://blog.hootsuite.com/reach-vs-impressions/

**General Notes**:

* Need Research Questions:
  + RQ1: Does the YouTube recommendation algorithm steer users towards extreme content?
  + RQ2: How frequent does YouTube provide recommendations outside of the original video’s ideology or tag?
  + RQ3: How accessible are other communities from the original community?
* BETTER TO GO WITH 2 AT THE UNDERGRAD LEVEL