

Using Network Analysis to Characterize Racist Parler Users

Pranav Venkatesh, Department of Computer Science, The University of Texas at Austin, Email: pranav.venkatesh@utexas.edu

Akaash Kolluri, Clements High School, Email: arkolluri@gmail.com

Nikhil Kolluri, Department of Electrical and Computer Engineering, The University of Texas at Austin, Email: nkolluri@utexas.edu

Dr. Dhiraj Murthy, School of Journalism and Media, Moody College of Communication, Email: Dhiraj.Murthy@austin.utexas.edu



Background

- Launched in September 2018, Parler became one of the most prominent US right-wing social media platforms.
- Parler had an ideologically polarized user base, where significant racist, misinformative, and violent content spread widely.
- Network-based methods have previously been successful at analyzing and tracking harmful content spread.

Research Questions

RQ1: Can we identify racist content on Parler at scale while simultaneously characterizing content at the user level?

RQ2: What groups is racist content on Parler targeting most?

RQ3: What can a network of user-level interactions tell us about groups on Parler?

Methods

Dataset Creation:

- Queried a large-scale Parler dataset [1] to find racist posts, users, and hashtags. Using a seed set of racist posts, “Best-First Searched” for more racist content by manually choosing the most racist content connected to the seed set and expanding to content connected to those new content.
- Added racist users, tokens, and hashtags to our data corpus.
- Identified 55 highly-racist users and compiled all their posts to make dataset ‘A’ (233K posts). Compiled all the replying posts to these users to create dataset ‘B’ (231K posts). Found the most frequent terms used by the users and queried all posts containing those terms to develop dataset ‘C’ (75K posts).

Cascade Analysis:

- Cascade connections occur when a user responds to a post, and both the post and the response contain the same racist content, determined by the presence of pre-defined racist terms.
- Cascades of racist terms and groups of racist content were mapped on dataset B; their size, breadth, and depth were measured.

Network-based Methods:

- Nodes (circles in Figure 1) are users, edges (lines in Figure 1) are mentions/replies/comments from original poster to replier.
- Based on categories manually assigned to hashtags, nodes are assigned colors based on hashtags used (see Figure 1).

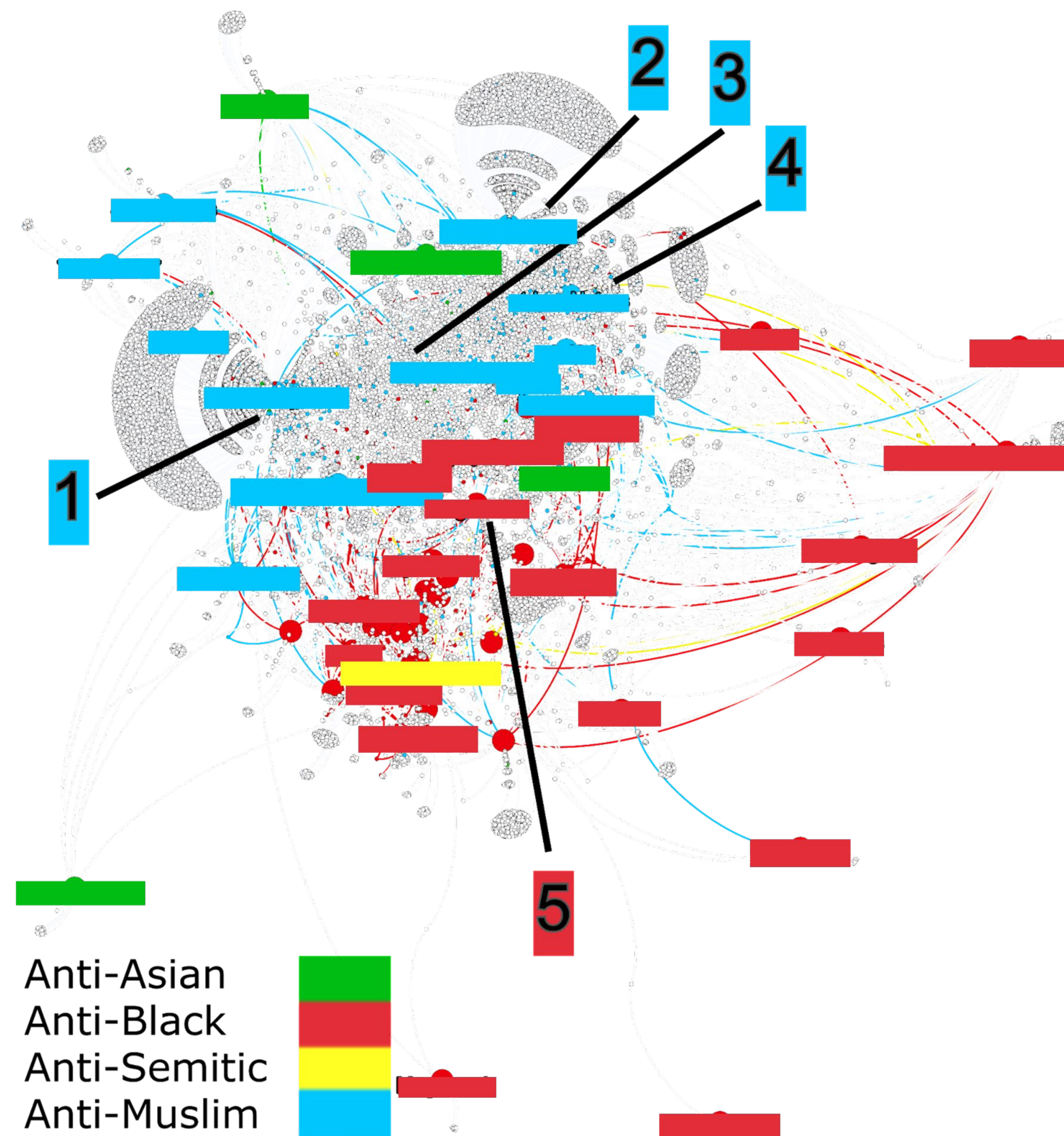


Figure 1: Graph of 55 highly racist **anonymized** users and repliers derived from datasets A and B. 1–5 are the most frequent posting users. White nodes/edges are unclassified users.

Term	Post Frequency	Cascade Frequency	Cascade to Post Percent	Average Cascade Depth	Average Cascade Size
n-word	11482	158	1.37%	2.01	2.3
whitelivesmatter	14160	86	0.61%	2.01	2.43
hitlerwasright	122	1	0.82%	2	2
chinesevirus	7734	9	0.12%	2.11	2.44
islamist	2873	15	0.52%	2	2.27

Table 1: Cascade attributes from 75k posts containing highly racist terms (dataset C)

Content	Post Frequency	Cascade Frequency	Cascade to Post Percent	Average Cascade Depth	Average Cascade Size
Anti-Black	25984	252	0.97%	2.01	2.34
Anti-Jewish	904	6	0.66%	2	2.33
Anti-Asian	26762	121	0.45%	2.03	2.29
Anti-Muslim	17548	89	0.51%	2	2.16

Table 2: Cascade attributes from 75k posts containing highly racist terms (dataset C)

Results

- Our “Best-First Search” method of using existing content to identify more content successfully found highly racist users (dataset A) and what terms are often used in racist posts (dataset C).
- Network analysis and visualization of content spread by users in dataset A and to their repliers in dataset B found that a small number of users/posts have an outsized impact on racist content production and consumption on Parler.
- Users with an outsized impact (based on eigenvector centrality) are predominantly Trump supporters, anti-vaccine, and white males who referenced beliefs such as QAnon and Nazism.
- Posts that are Anti-Black or Anti-Muslim were most frequent (measured by centrality scores).
- Anti-Black content was found to be 1.5 to 2 times more likely than other forms of racism to inspire similar racism in future posts.

Conclusion

- We successfully characterized the diffusion of highly racist content from user to user on Parler.
- We built a robust framework to generate network visualizations of specified Parler users for any time-period, assisting in highlighting the nuanced interactions and saturation of racist sub-groups.
- We contributed analysis suggesting that Anti-Black content had the highest risk of spread when compared to other forms of racism.
- Some of the most influential users on the network promoted misinformation, conspiracy theory, and racist ideology.
- Moderating this racist content of a small number of users could help platforms mitigate harm on social media platforms.

Acknowledgments

Thanks to Bin Chen, Torie Kim, and Jo Lukito for their feedback.

References

- [1] Aliapoulos, M., Bevensee, E., Blackburn, J., Bradlyn, B., De Cristofaro, E., Stringhini, G., & Zannettou S. (2021). A Large Open Dataset from the Parler Social Network (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.4442460>