# Characterizing Twitter users who engage in Adversarial Interactions against Political Candidates

**Yiqing Hua** Mor Naaman Thomas Ristenpart

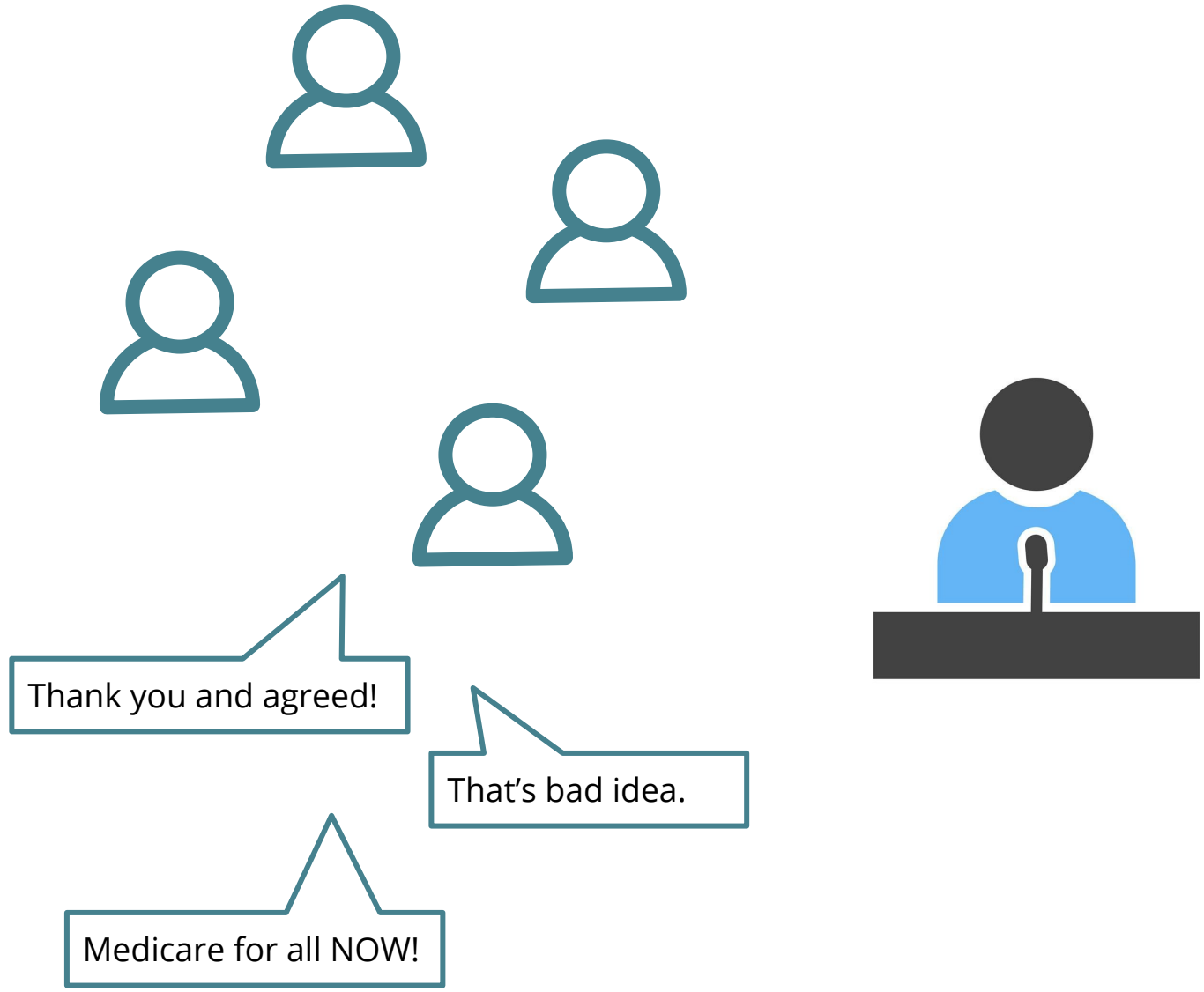
Cornell Tech, Cornell University

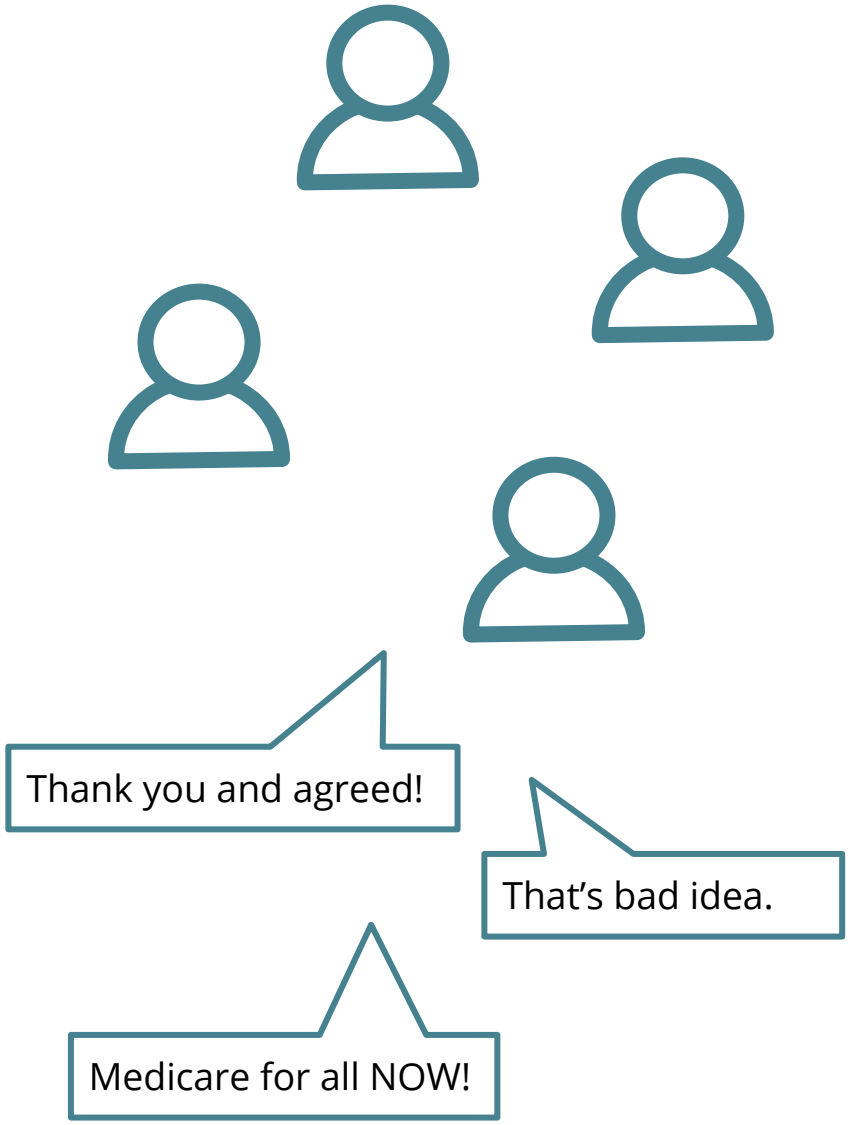
Presentation for CHI 2020
Find the slides and paper on yiqing-hua.com

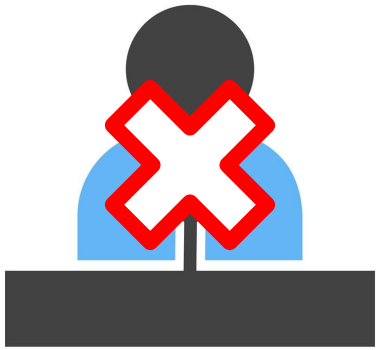
CORNELL TECH
HOME OF THE JACOBS INSTITUTE

- Connect with constituents
- Express opinions
- Campaign for the race

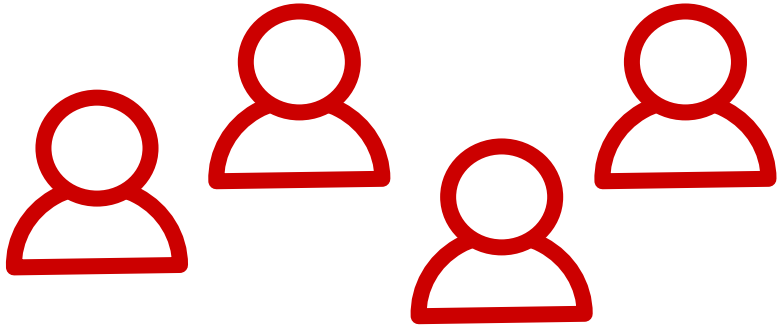
CORNELL TECH
HOME OF THE JACOBS INSTITUTE


Thank you and agreed!
That’s bad idea.
Medicare for all NOW!


Thank you and agreed!
That's bad idea.
Medicare for all NOW!




SHUT UP!
Most annoying woman ever…
Alexandria Occasionally-Coherent!

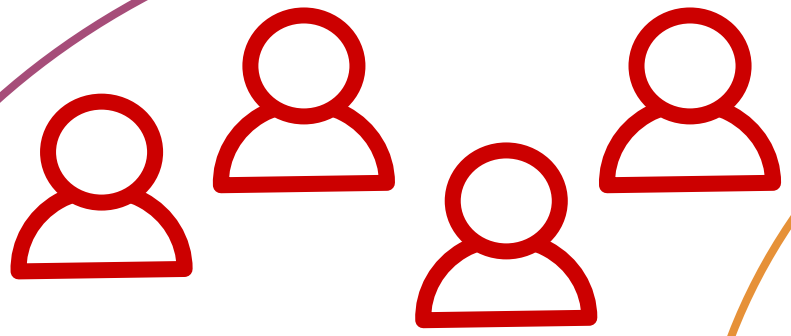


**10%** of the users created over **35%** of the adversarial interactions.

Adversarial users exhibit **different** behavioral patterns than normal user, showing a tendency to **seek out conflicts**.

They involve in **fewer supportive** interactions and pay more attention to **opponent** candidates.

CORNELL TECH
HOME OF THE JACOBS INSTITUTE



Chatzakou et al. (2017)
ElSherif et al. (2018)
Ribeiro et al. (2018)

Characterizing Adversarial Users

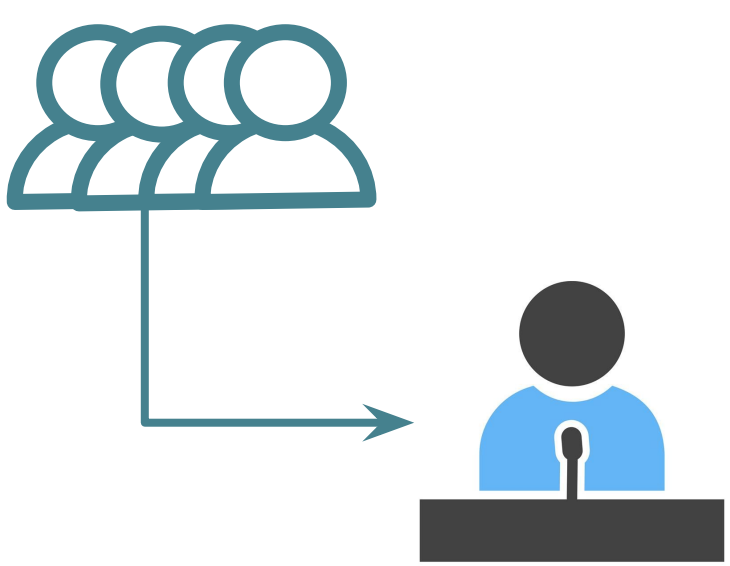
**Adversarial users in political context**

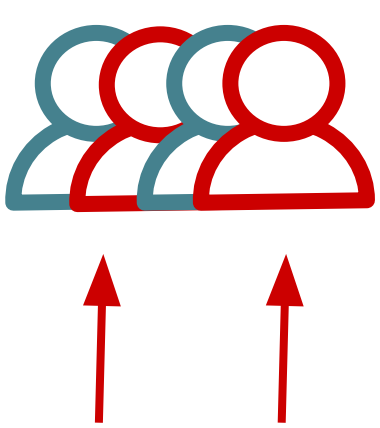
Gorrell et al. (2018)
Theocharis et al. (2020)

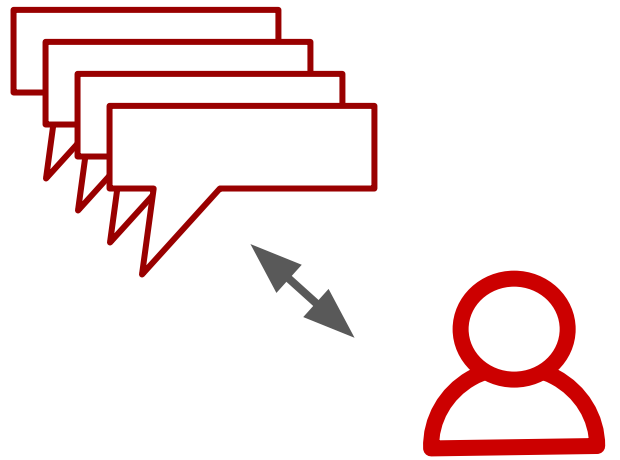
*Hua et al. (ICWSM 2020)*
*check it out on yiqing-hua.com*

Adversarial Interactions in political context

CORNELL TECH
HOME OF THE JACOBS INSTITUTE







Data Collection

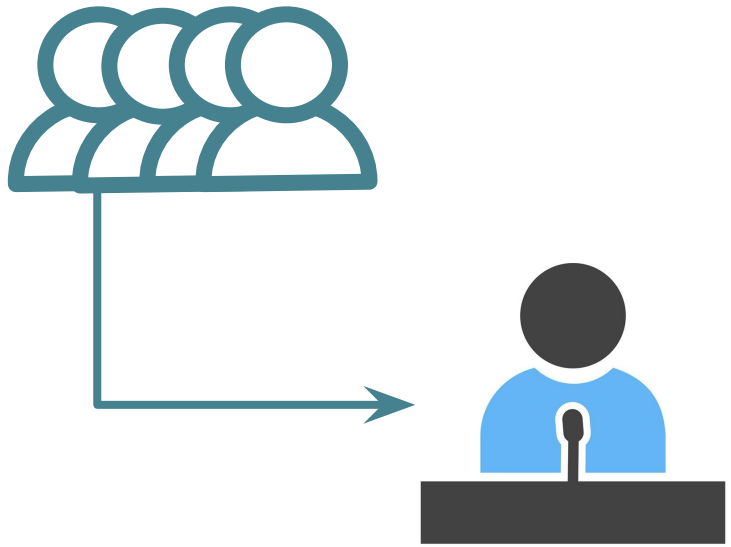
Identify Adversarial Interactions

Correlate *User Characteristics* with *Amount of Adversarial Interactions*

CORNELL TECH
HOME OF THE JACOBS INSTITUTE

# U.S. Midterm Election Twitter Dataset 2018

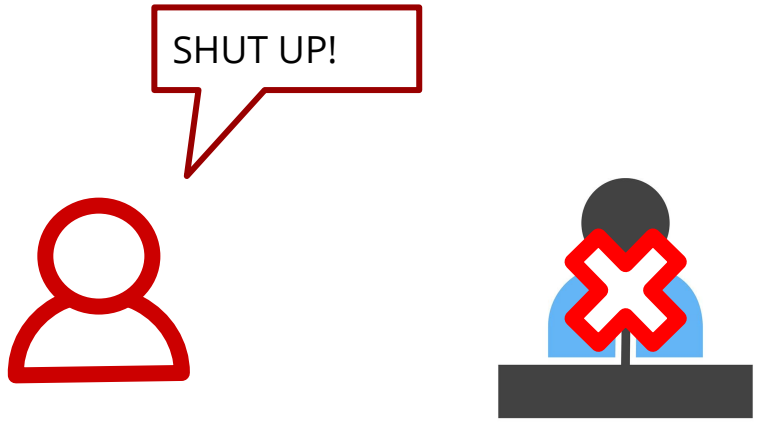
**1.2M** user replies to **786** candidates running for U.S. House of Representatives (87%) between **September 17th, 2018** to **November 6th** from **0.4M** users



Dataset published on Figshare
Find the link at yiqing-hua.com

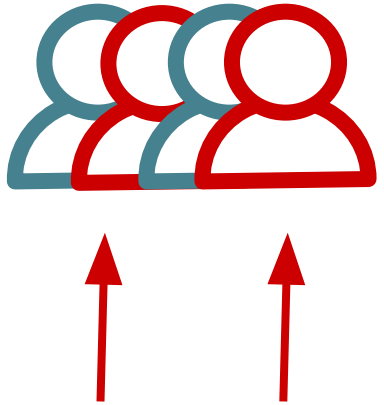
CORNELL TECH
HOME OF THE JACOBS INSTITUTE

# Adversarial Interactions




Behaviors on social media that intended to **hurt, embarrass, or humiliate** a targeted individual.

# Identify Adversarial Interactions

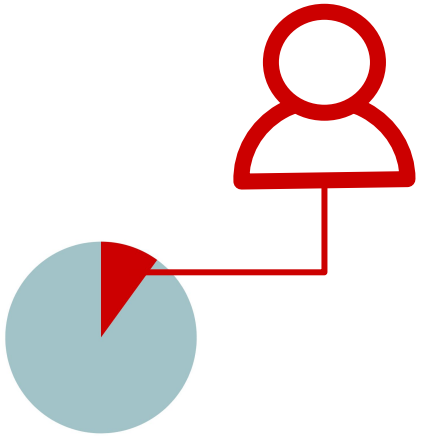


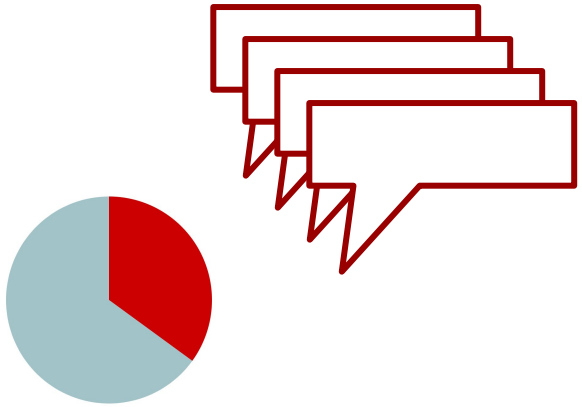
Use **Toxicity** scoring from Perspective API to identify adversarial interactions

*"a rude, disrespectful, or unreasonable comment that is likely to make you leave a discussion."*

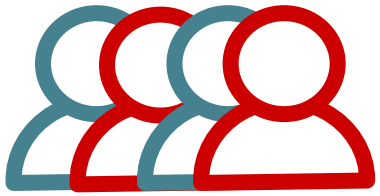
Please refer to the details regarding validating this approach in our paper.

CORNELL TECH
HOME OF THE JACOBS INSTITUTE





**10%** of the users created over **35%** of the adversarial replies.



# Moderately active users

posted more than **3**, no more than **30** interactions

**21%** of all users, contributed **50%** of all interactions and **52%** of the adversarial interactions.

CORNELL TECH
HOME OF THE JACOBS INSTITUTE

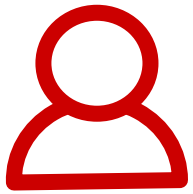
# Correlate *User Characteristics* with *Amount of Adversarial Interactions*

**Control**

Replies to Candidates

**Engagement in Political Activities**

Supportive interactions with candidates
Centrality in politically engaged crowd
Attention to opponent candidates
Partisan-ness in profile



**Basic User Features**

Number of Followers
Number of Days on Twitter
Verified on Twitter (to approximate anonymity)

**Adversarial Activities by Twitter Friends**

CORNELL TECH
HOME OF THE JACOBS INSTITUTE

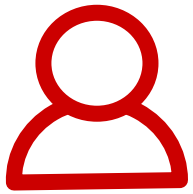
# Correlate *User Characteristics* with *Amount of Adversarial Interactions*

**Control**

**+ Replies to Candidates**

**Engagement in Political Activities**

- **Supportive interactions with candidates**
- **Centrality in politically engaged crowd**
+ **Attention to opponent candidates**
+ **Partisan-ness in profile**



**Basic User Features**

- **Number of Followers**

Number of Days on Twitter
Verified on Twitter (to approximate anonymity)

$$R^2 = 0.231$$

- **Adversarial Activities by Twitter Friends**

Adversarial users exhibit **different** behavioral patterns than normal user.

# Correlate *User Characteristics* with *Amount of Adversarial Interactions*

**Control**

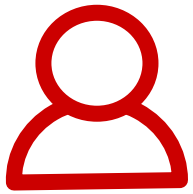
\+ **Replies to Candidates**

**Engagement in Political Activities**

- **Supportive interactions with candidates**
- **Centrality in politically engaged crowd**

\+ **Attention to opponent candidates**

\+ **Partisan-ness in profile**



**Basic User Features**

- **Number of Followers**

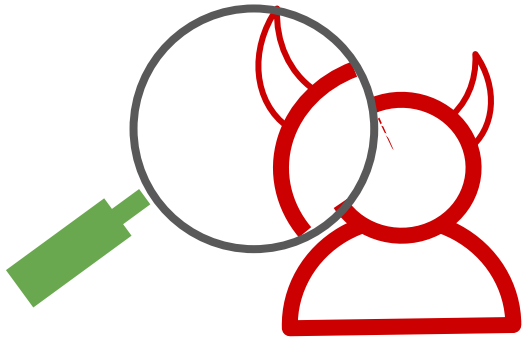
Number of Days on Twitter

Verified on Twitter (to approximate anonymity)

\+ **Adversarial Activities by Twitter Friends**

## Tendency to seek out conflicts

CORNELL TECH
HOME OF THE JACOBS INSTITUTE



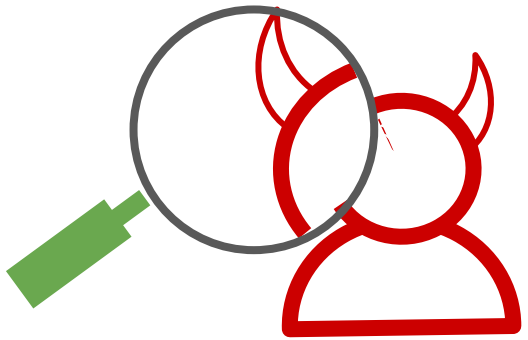
- **Supportive interactions with candidates**

  Measured using number of retweets and following

+ **Attention to opponent candidates**

  Measured using number of replies to opponent candidates

## What about the content of the interactions?

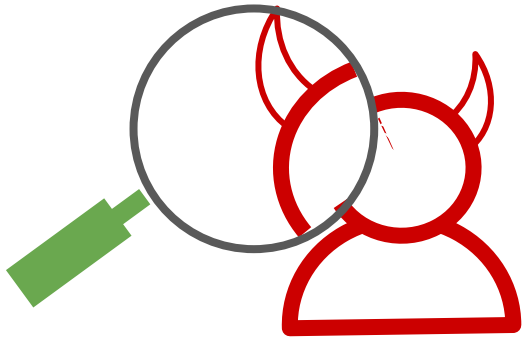


- **Supportive interactions with candidates**

  Do adversarial users post fewer supportive replies?

+ **Attention to opponent candidates**

  Are adversarial users more negative in their replies to candidates?



## Highly adversarial users

posted more than 10 adversarial interactions

**0.3%** of all the users.

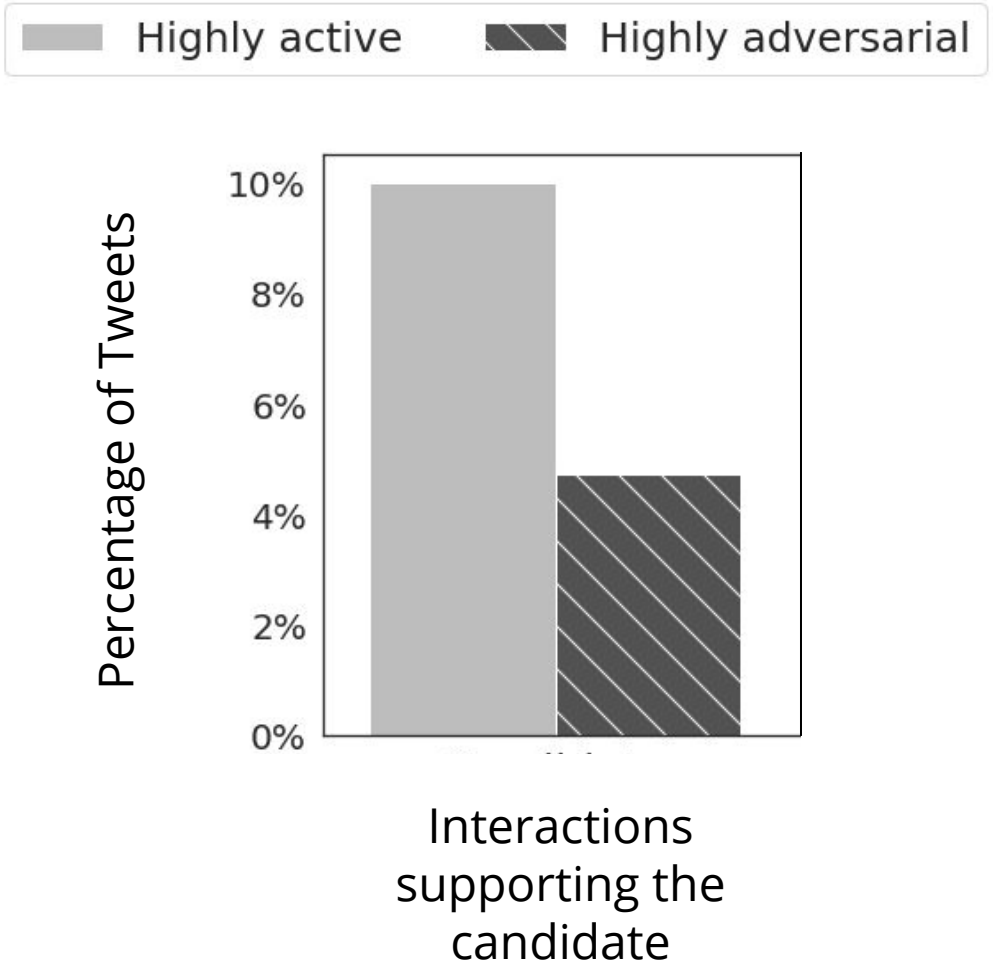
contributed **10%** of all adversarial interactions and **5.6%** of all interactions.

## Highly active users

posted more than 10 interactions

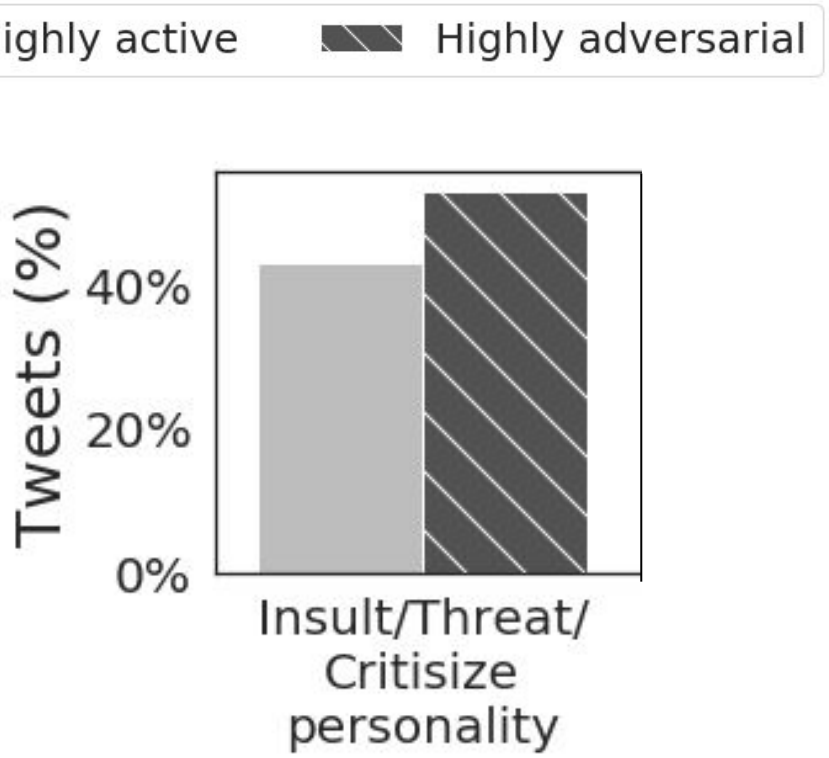
Randomly sample 200 adversarial and 200 non-adversarial interactions from each group. Perform manual labeling on the samples.

# Supportive Interactions

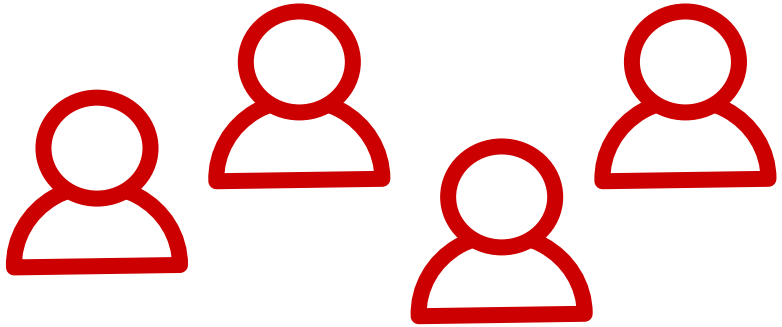



**Fewer** interactions **supporting** candidates themselves.

# Negative Interactions




**More negative** interactions at personal level.



**10%** of the users created over **35%** of the adversarial interactions.

Adversarial users exhibit **different** behavioral patterns than normal user, showing a tendency to **seek out conflicts**.

They involve in **fewer supportive** interactions and pay more attention to **opponent** candidates.

CORNELL TECH
HOME OF THE JACOBS INSTITUTE

# Adversarial interactions with political candidates

Characterizing Twitter Users Who Engage in Adversarial Interactions against Political Candidates. [CHI2020]

Towards Measuring Adversarial Twitter Interactions against Candidates in the US Midterm Elections. [ICWSM2020]

yiqing-hua.com

yiqing@cs.cornell.edu