#BidenCheated Hashtag and How Quickly It Spread on Twitter

Second Forensics Study CS895 F2020

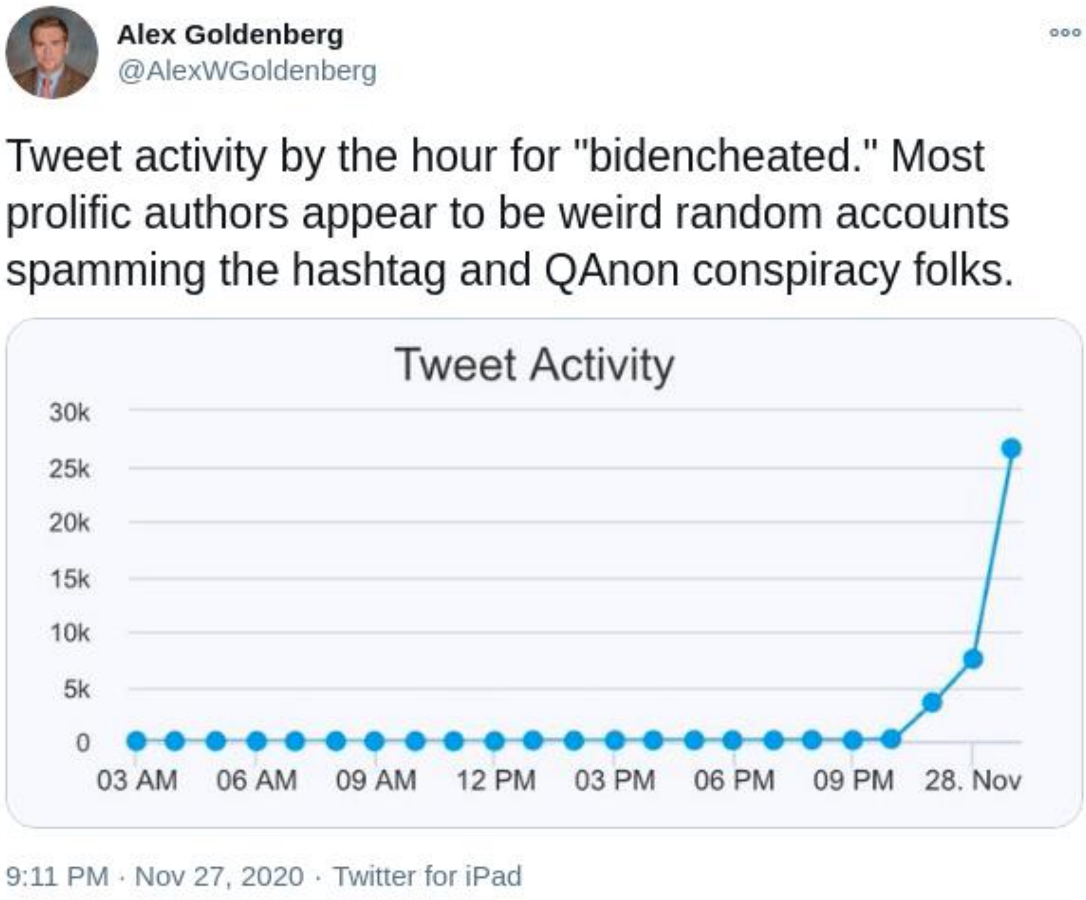
2020-12-17

Peter Foytik

At end of November 2020 Twitter accounts were seen spamming #BidenCheated

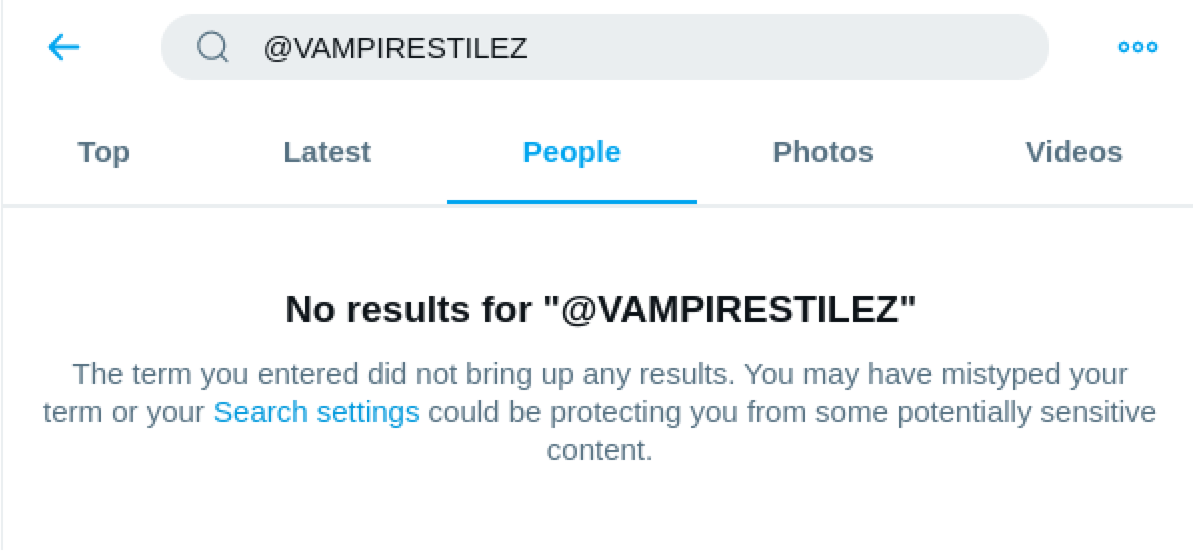
- How far did these spammed tweets reach across a potential twitter network?
    - The tweets are public but how many accounts had direct access to it (through followers)
    - How many of these following accounts also participated in the spamming act
- What was the rate of the spread of information based on follower's activity?
- Was there bot activity that is detectable and how much did this contribute to the spread of disinformation?

Alex Goldenberg tweeted November 27 on how many users were tweeting #BidenCheated.



https://twitter.com/AlexWGoldenberg/status/1332507392840503298

Two accounts mentioned from Alex Goldenberg "Starcisco" and "Vampirestilez"



Managed to scrape tweets with "#BidenCheated" and followers for VAMPIRESTILEZ before account was removed!

Unfortunately, the account was not archived

Things that likely contribute to the spread of information

- Network reach
  - Number of followers
  - Willingness of followers to contribute
- Information Frequency
  - Speed at which the information is transmitted
  - Is there a sweet spot?
    - Too fast might give away malicious motive
    - Too slow might not get the message out

https://blog.twitter.com/engineering/en_us/topics/insights/2017/using-deep-learning-at-scale-in-twitters-timelines.html

Tools used to obtain data for analysis

- Twint (https://github.com/twintproject/twint)
  - scrape tweets containing "#BidenCheated"

```
twint –s "#BidenCheated --since 2020-10-28
```

- Twitter API (Twitter Developer https://developer.twitter.com/en)
  - Scrape follower accounts of participants
- Botometer API (https://botometer.osome.iu.edu/api RapidAPI)
  - Score participants to get an idea of the possibility of being a bot
- Python

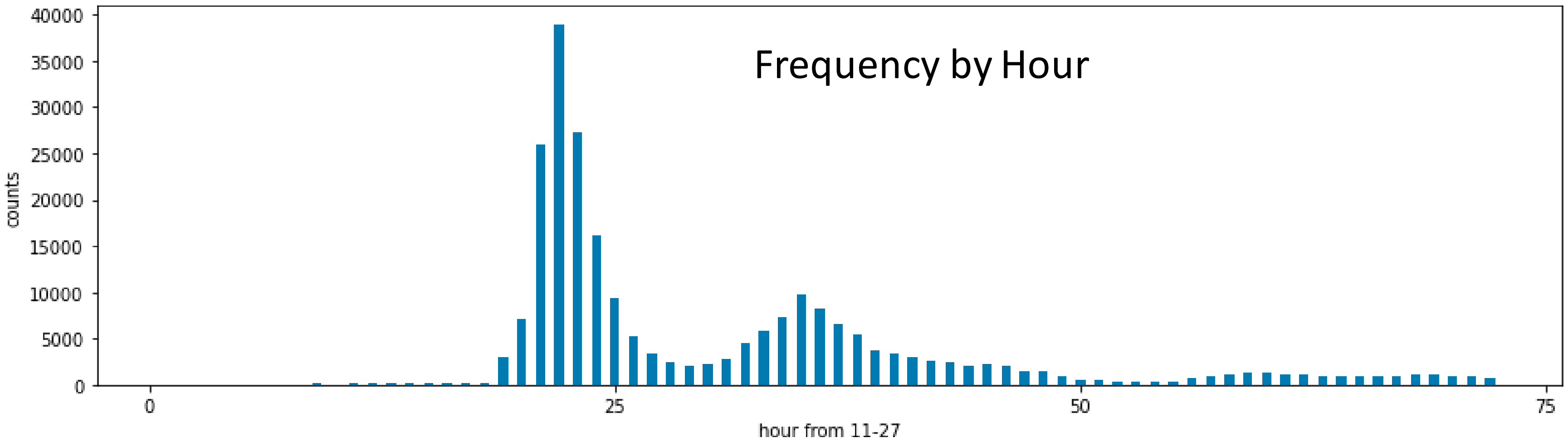
Summary #BidenCheated Tweets

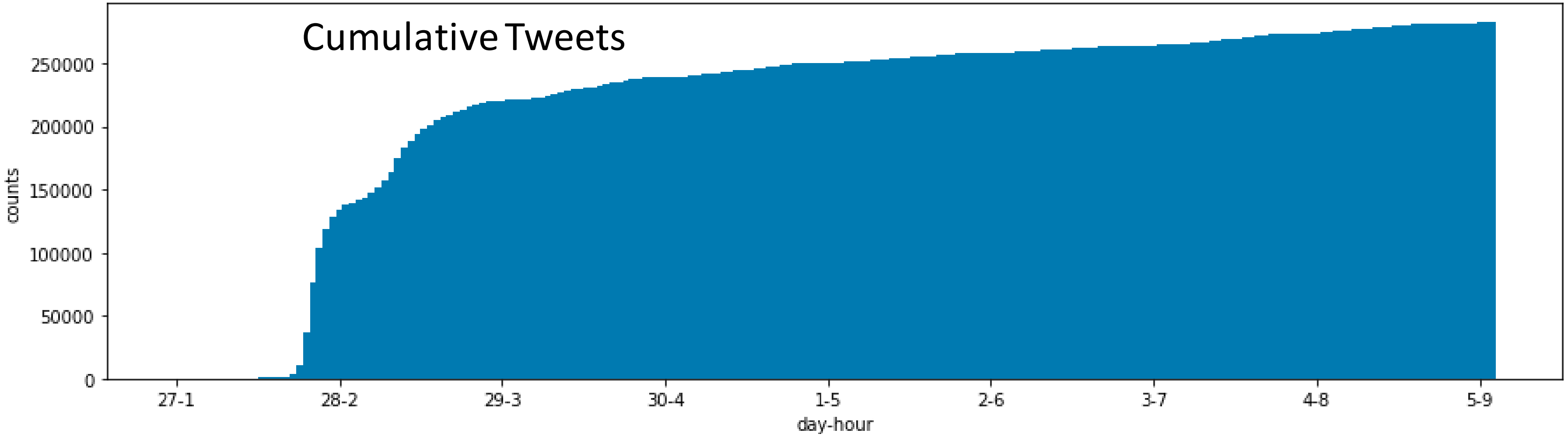
- Total data set covers dates from 2008-10-25 – 2020-12-05
- Total number of tweets scraped: 305,781
- Unique user accounts: 80,405
- Number of accounts follwers scraped: 366
- Total accounts analyzed: 630,794
- Total number scored with Botometer: 6,456

Rate of tweets with BidenCheated hashtag

Max tweets per hour (TPH):
- Nov 27
- hour 22
- TPH = 38,991

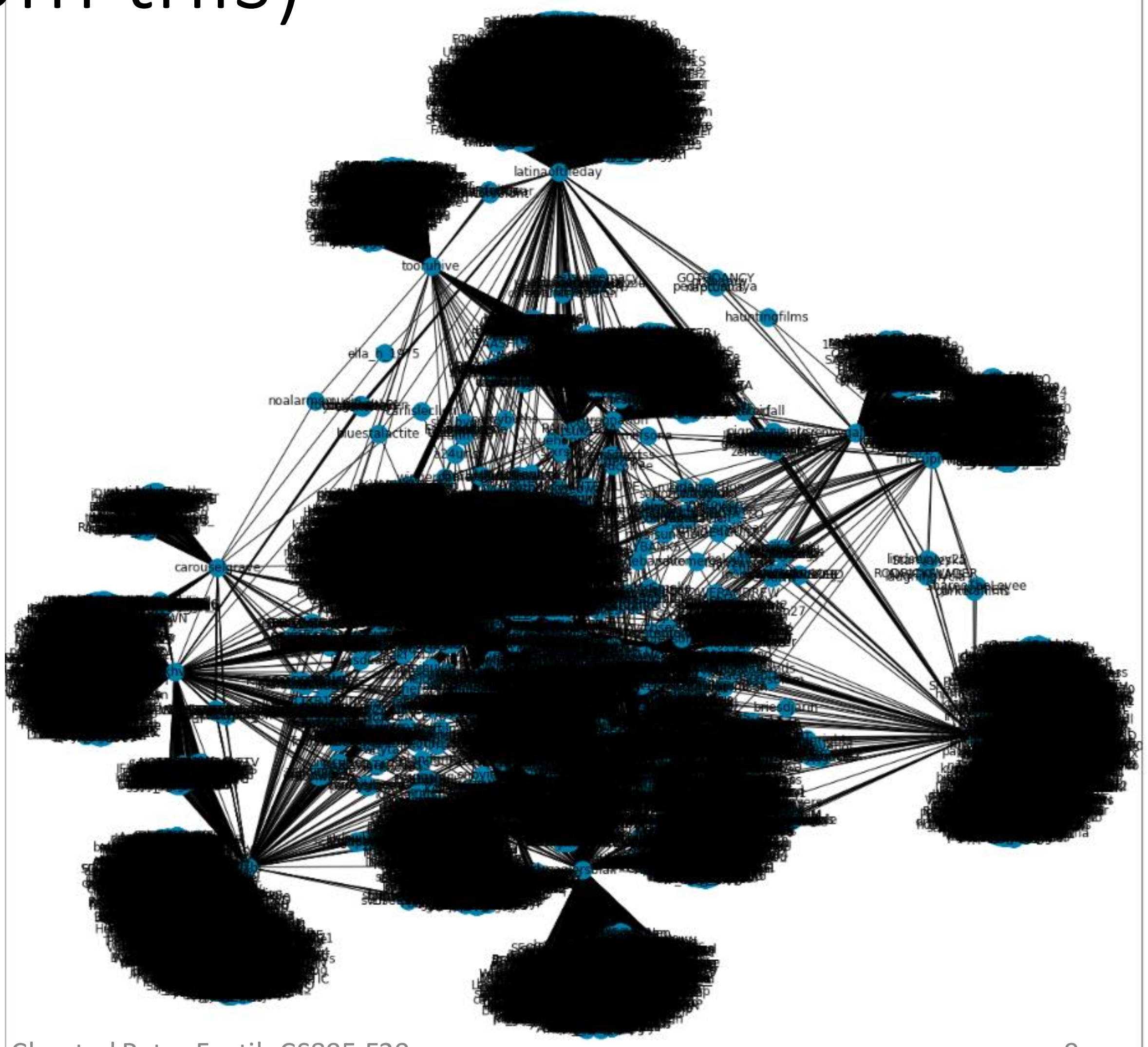
Average TPH:
- 1,395





Full network of all participants and their followers (Can't really get much from this)

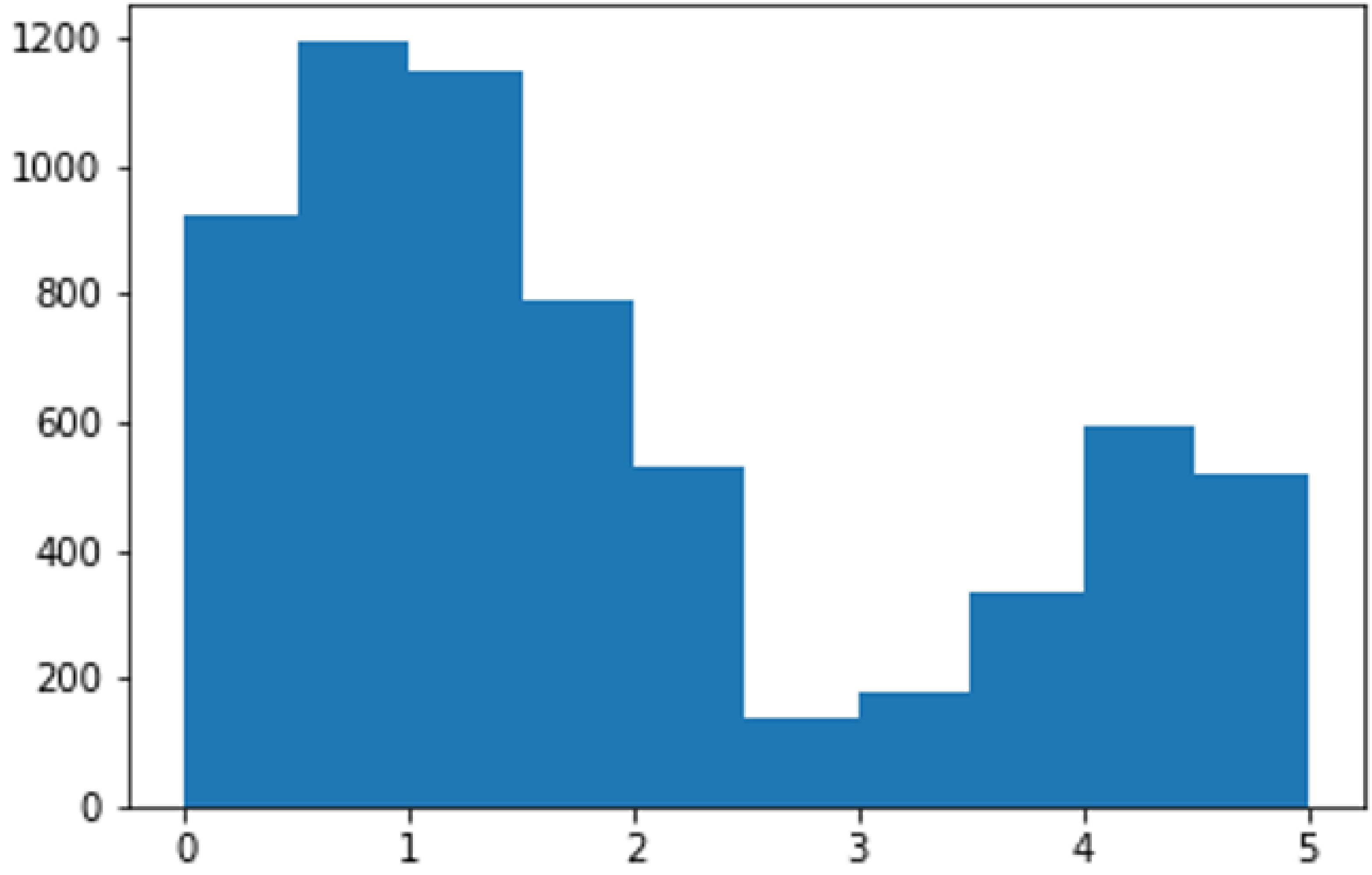
- Parent accounts: 366
  - All parent accounts were participants in BidenCheated hashtag
- Average follower per account: 1,884
- Max follower per account: 138,511



Batometer results for the collection involved in the BidenCheated spam

- Botometer scores are retrieved for all accounts that participated in BidenCheated
- Total accounts retrieved
  - 6,938
- Average Bot Score:
  - 1.9
- Median Bot Score:
  - 1.4
- Number of Bots >= 4
  - 1,111

Histogram of total bot scores



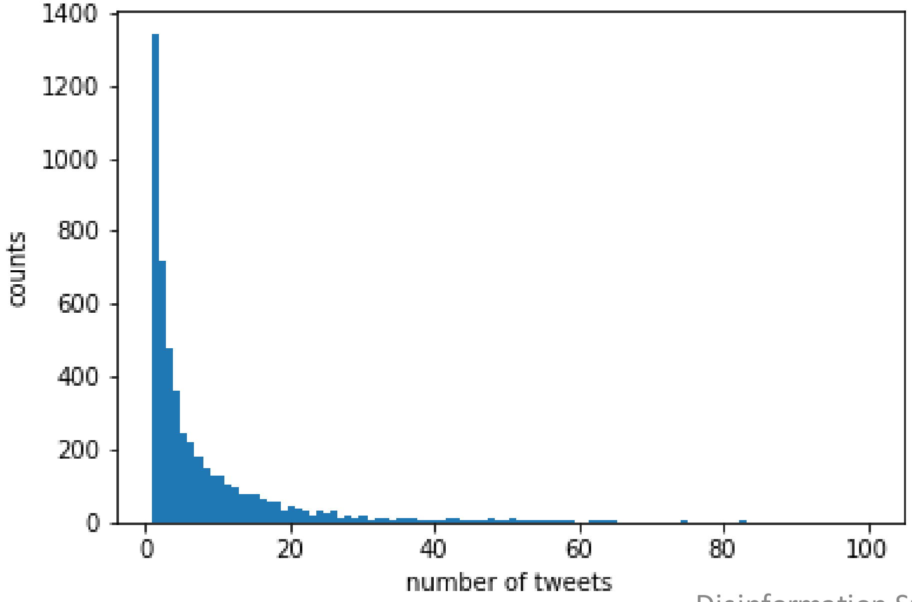
Bot contribution to BidenCheated (Botometer score >= 4)

Total Unique accounts = 102,979

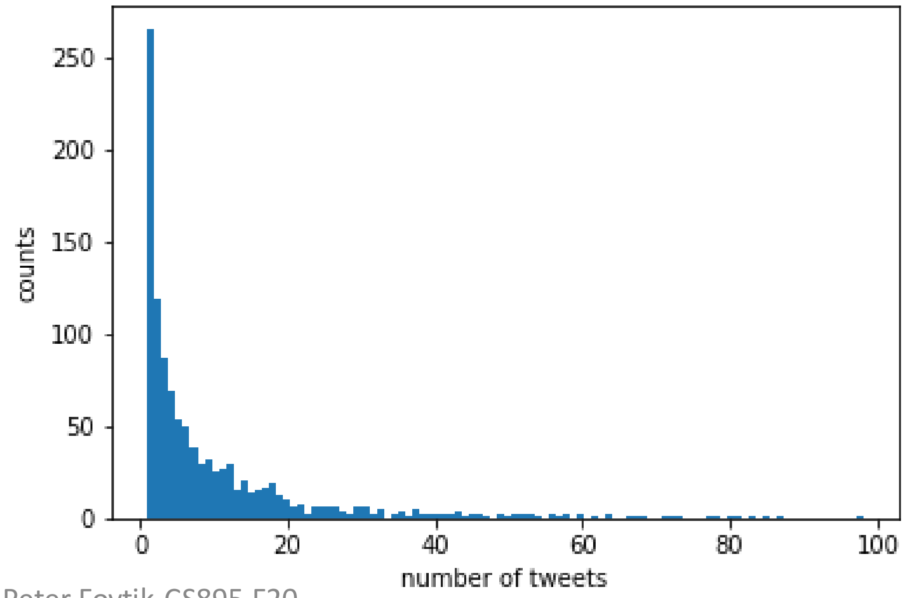
Total **non**Bot accounts = 5,260

Total Bot accounts = 1,111

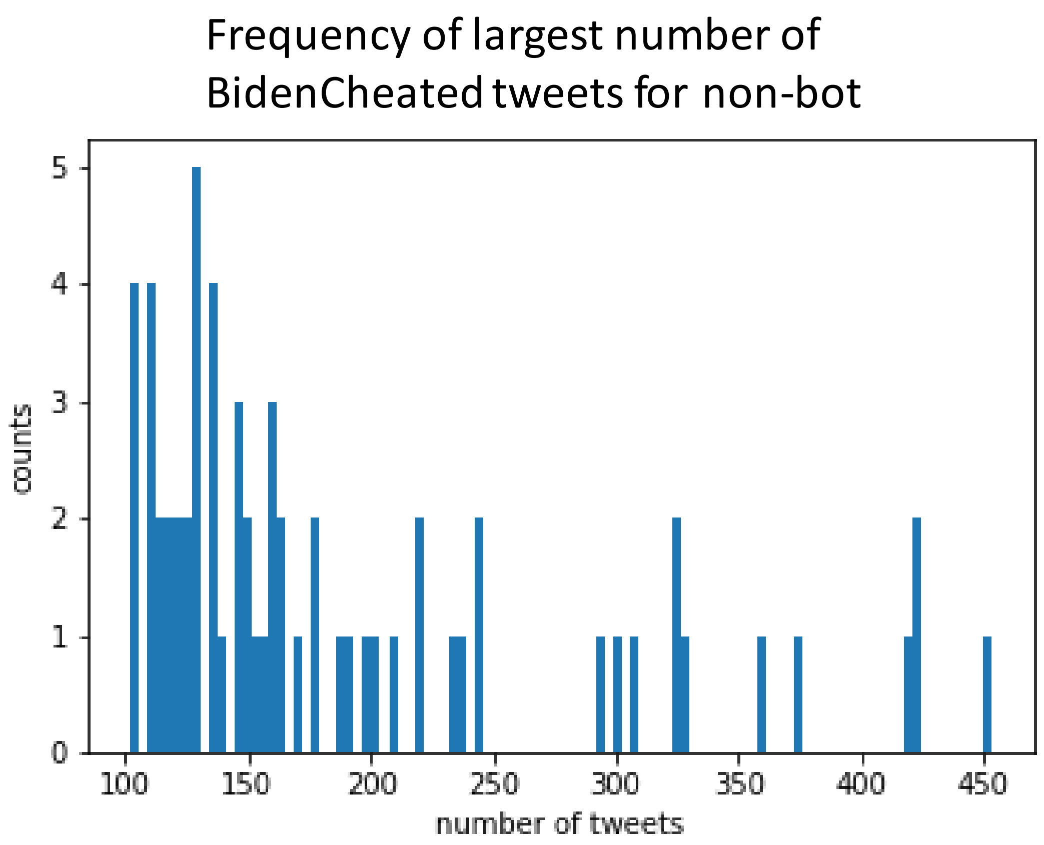
Frequency of number of BidenCheated tweets for **non**Bot

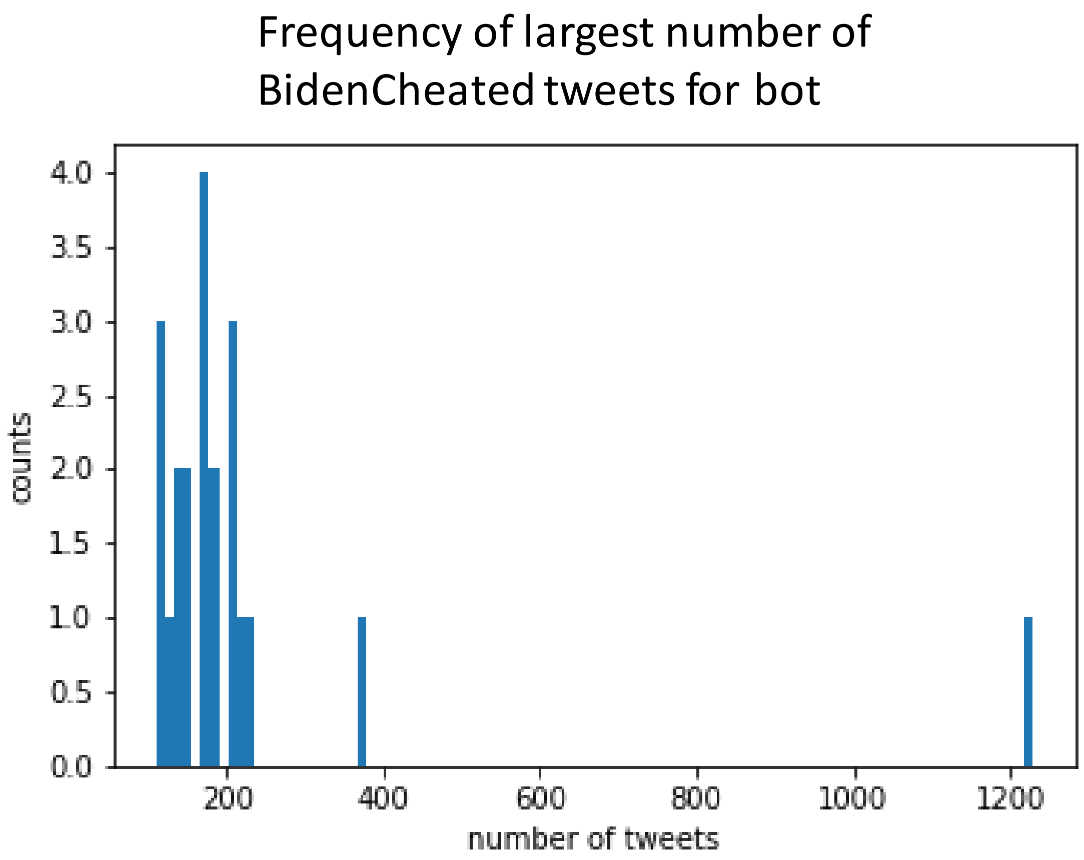


Frequency of number of BidenCheated tweets for Bot



Maximum frequency of tweets from non-bot users and bot users

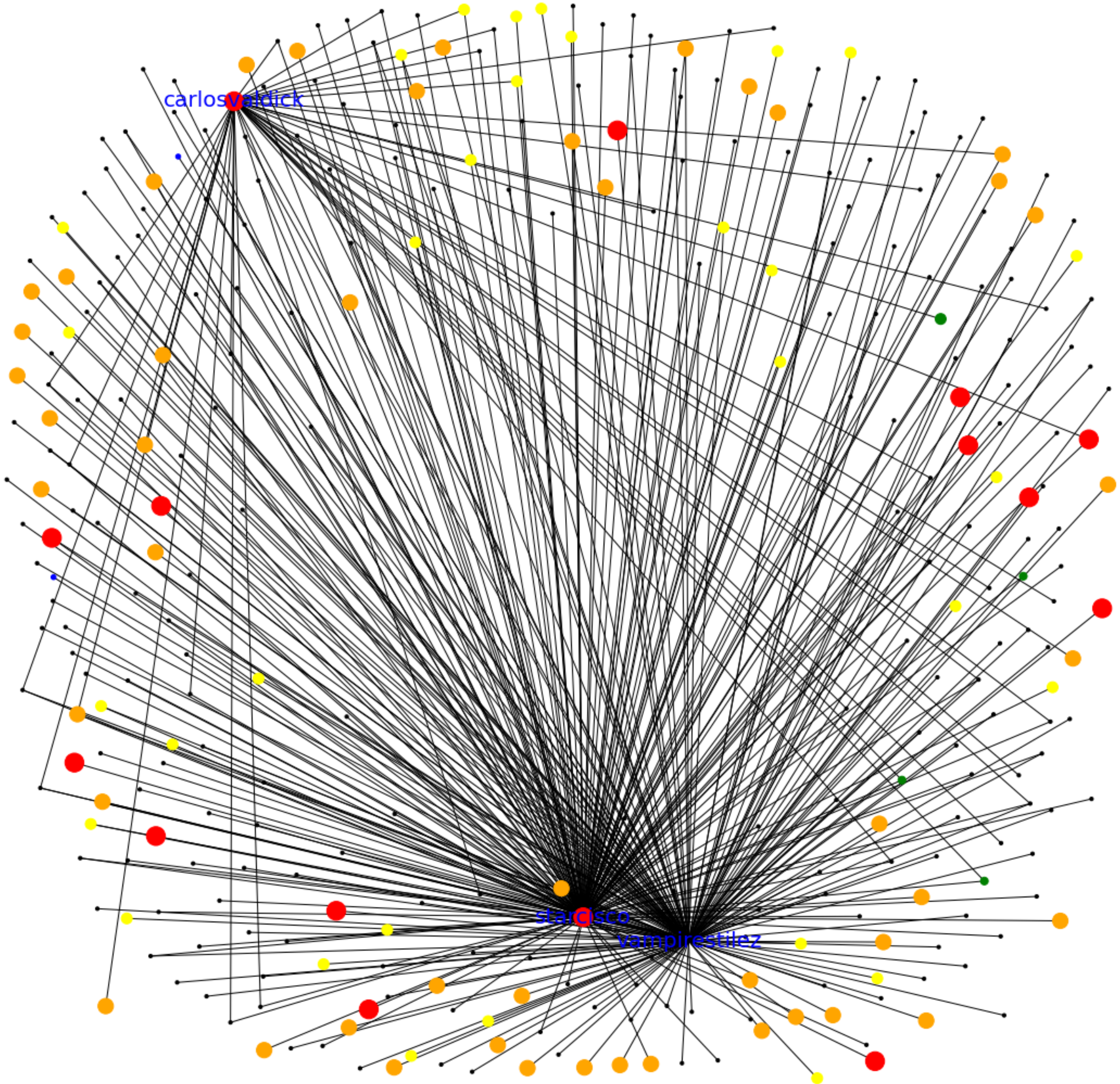




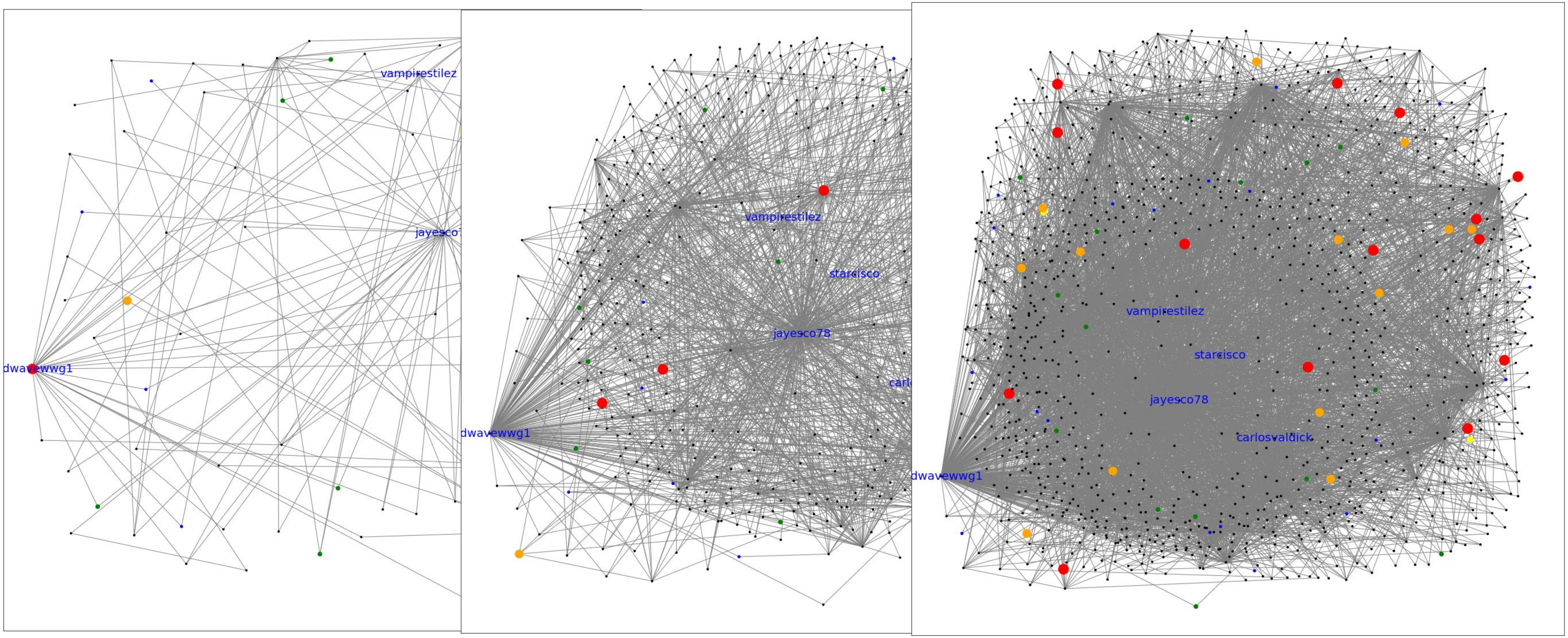
Three targets: carlosvaldick, starcisco, vampirstilez

Nodes are colored based on bot score (0-5 greater number more bot like).

- 0
- 1
- 2
- 3
- 4
- 5



BidenCheated propagation over time and identified additional highly connected accounts

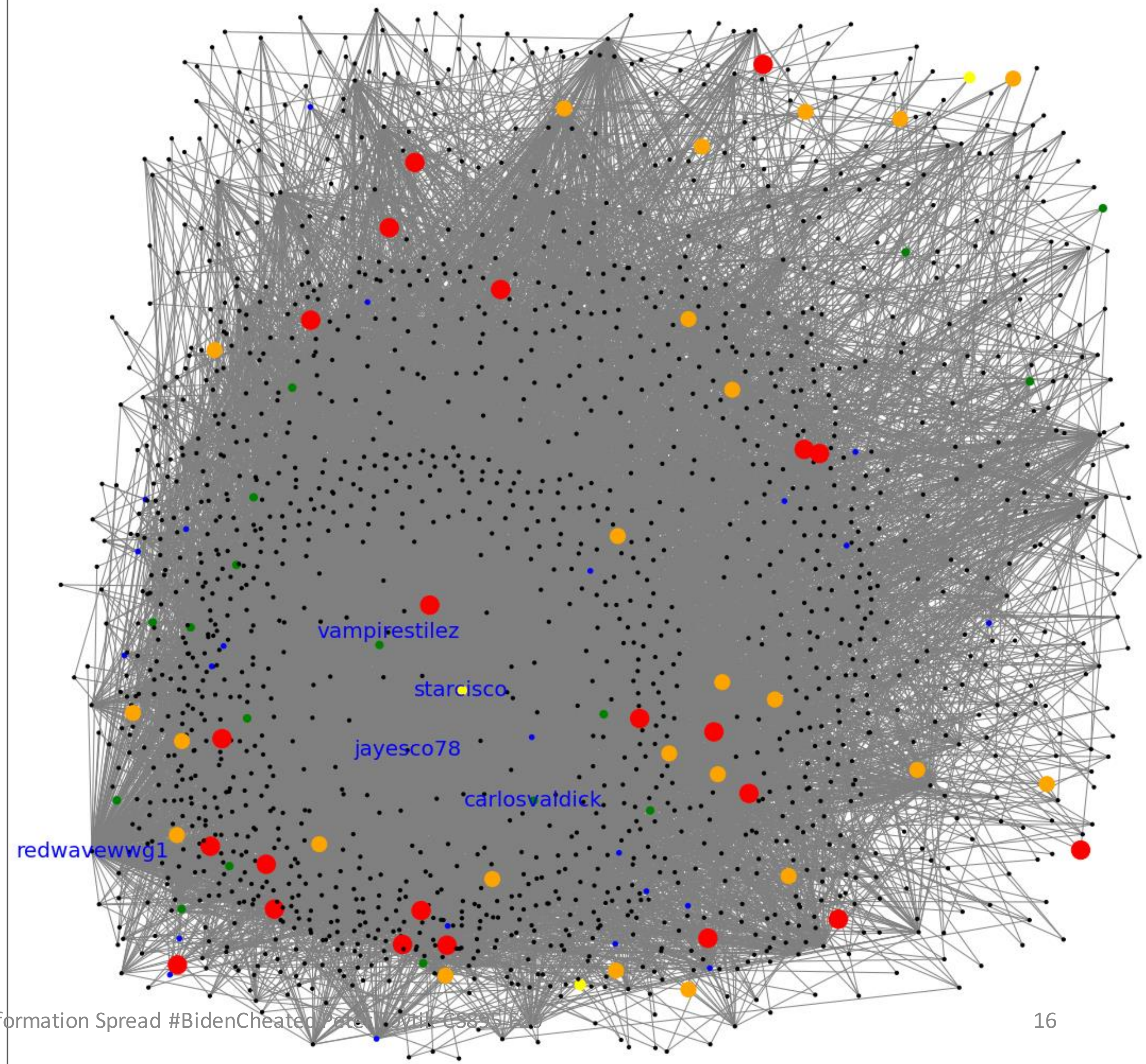


Final state at peak TPH

Average centrality = 0.005
Maximum Centrality = 0.29

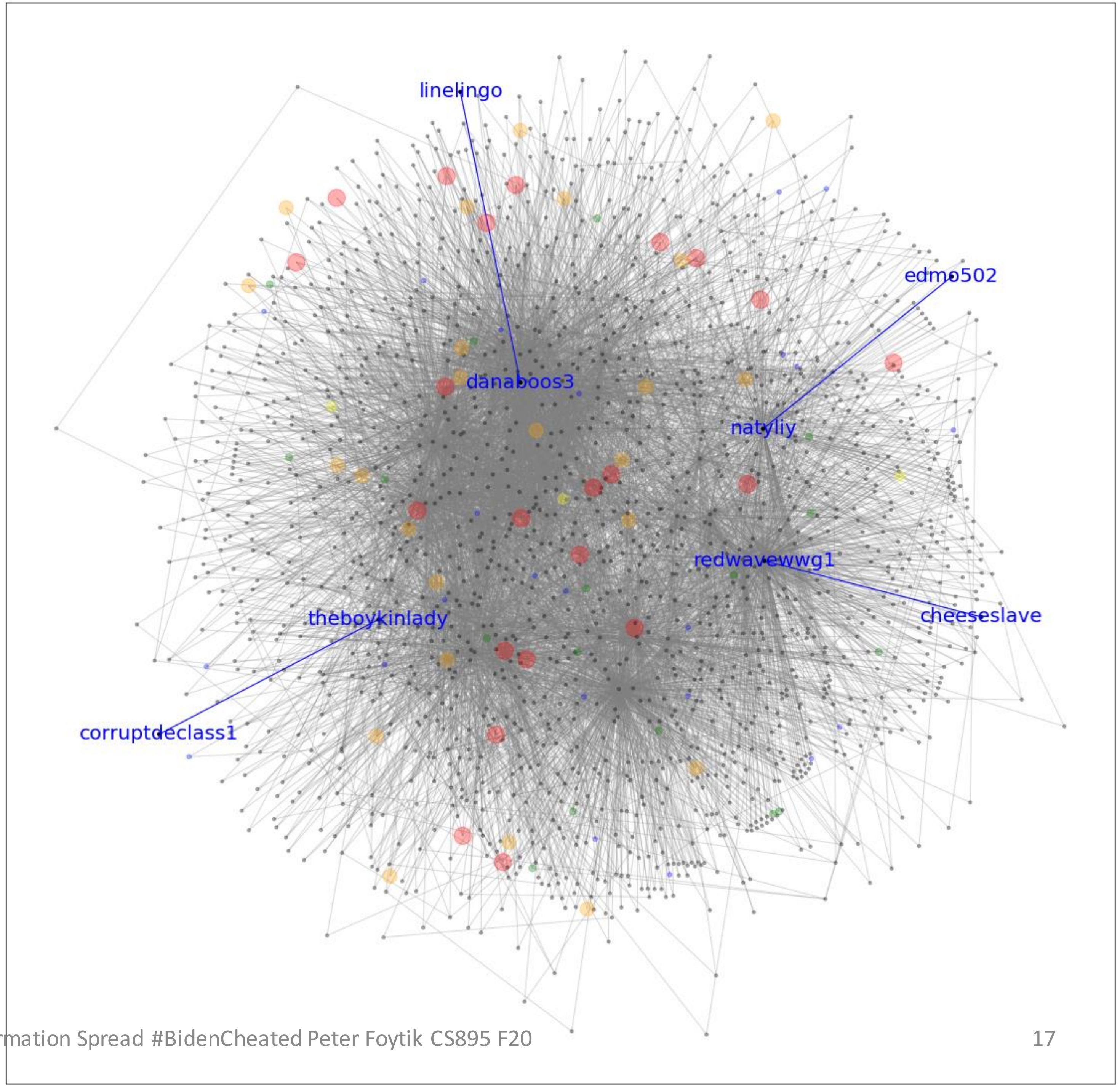
Average closeness = 0.317
Maximum Closeness = 0.45

Average betweeness = 0.001
Maximum betweenness = 0.246



4 bridge links were measured as critical links, contributing to the structure of the network

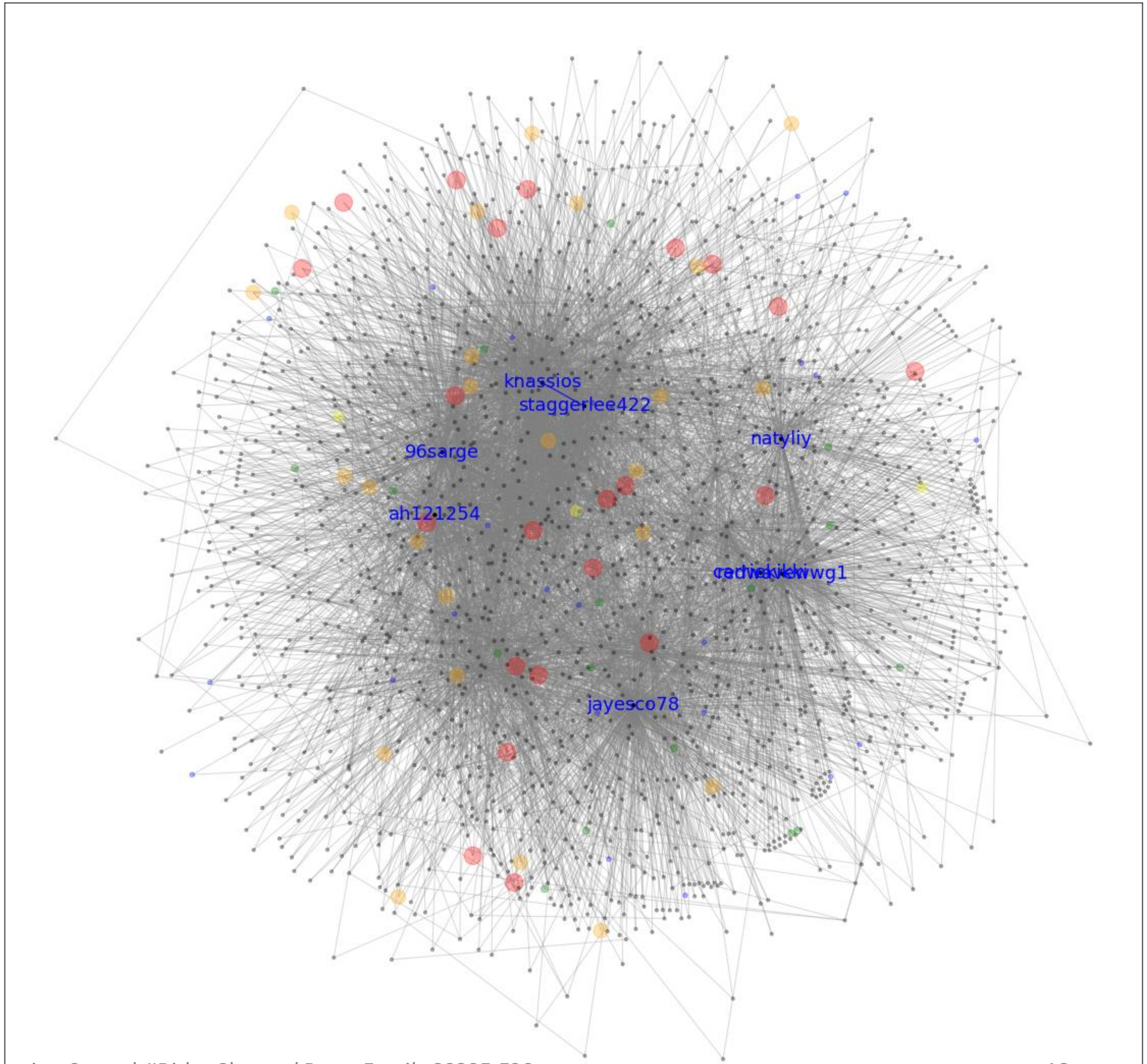
All nodes were nonBot



Network of high centrality nodes

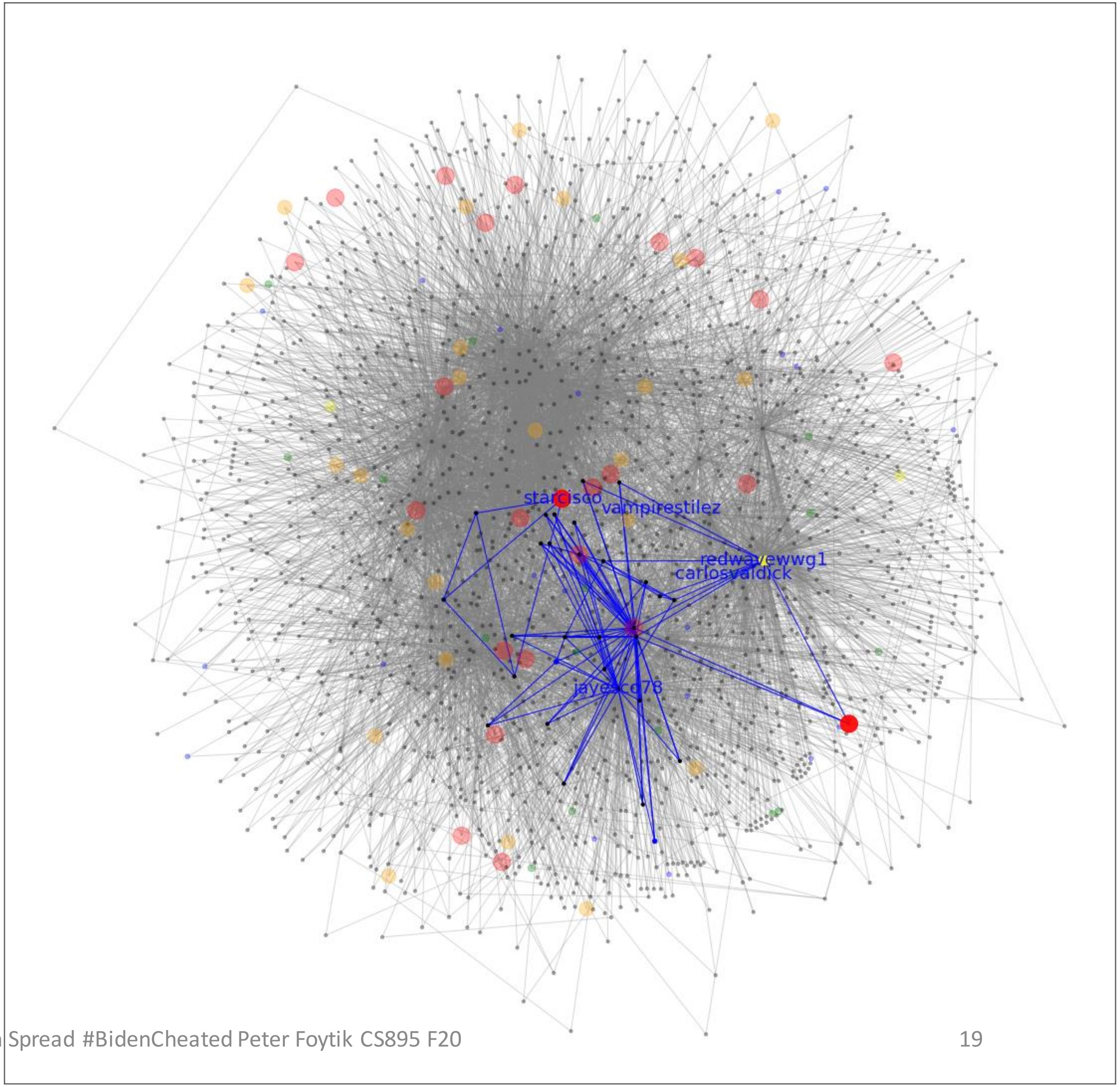
Blue nodes are highest degree centrality (most influence of followers)

Also all nonBot



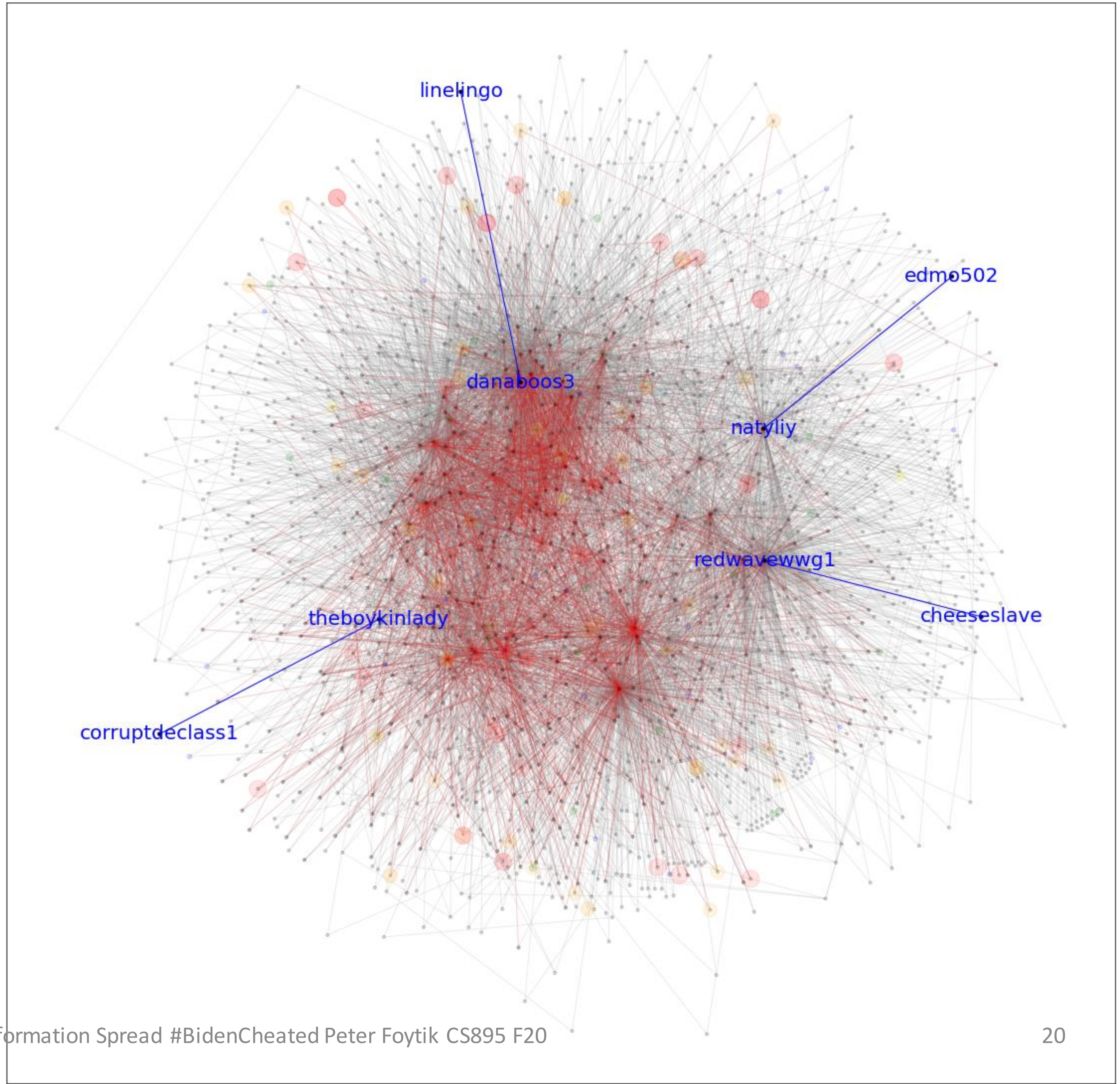
27 Cliques are derived of size 4 user accounts each containing targets listed by Alex Goldenberg

2 bots included in cliques



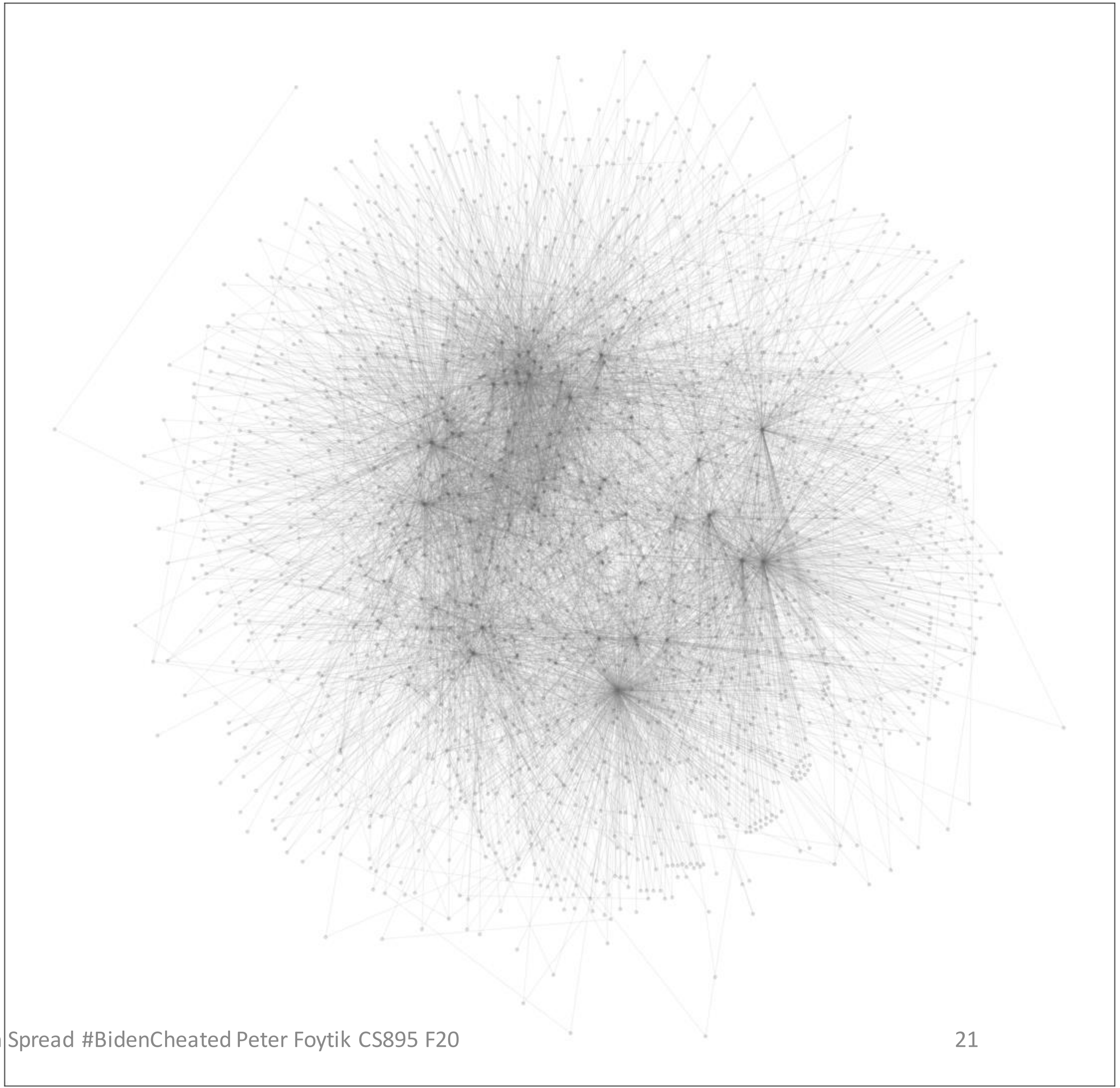
Bots overlayed with the bridge nodes and links

Areas of bots connect to ends of bridges



Removing bot nodes and measuring the difference in the network

- Bridges increased from 4 to 60
- Average centrality reduced 0.001
- Average closeness reduced slightly 0.01
- Average betweenness reduced slightly 0.0003



Major takeaways and conclusions

- Based on connectivity and rate of tweeting BidenCheated, bots (scored by botometer) did not appear to have as much influence in this scenario.
  - The three targets were not scored as bots
  - The main connected nodes were not bots
- Though bots contribute to the spread of disinformation (especially with higher rate of tweets), the network observed in this scenario relied on bridge connections of users that did not appear to be bots.
  - Bots are likely effective when used in conjunction with the non-bot accounts
- Results confirm findings from literature

Vosoughi, Soroush, Deb Roy, and Sinan Aral. "The spread of true and false news online." *Science* 359.6380 (2018): 1146-1151.

Future work with this dataset and measurements

- This data can contribute to a collection of scenarios that can represent disinformation spread based on environmental variables:
  - Bot influence
  - Network topology
  - Rate of tweet
- From observed scenarios models can be developed that represent the spread of disinformation based on the environment
- These models can run in a simulated environment to better understand and forecast rate of disinformation spread based on the environment