

Monica Powell
Data Visualization
March 3rd, 2016
Project URL: <http://aboutmonica.com/QMSS/forcedirected/>

Force Layout Visualization of the Twitter Mention Network for Bernie Sanders, Donald Trump and Hillary Clinton, Ted Cruz

The force-directed graph of the Twitter mention network of all four candidates shows that some nodes (Twitter users) have relatively high betweenness and centrality to the network overall. The Twitter mention network was constructed by combining the nodes and edges of the .csv data file that was generated from a 5 hour stream of Tweets that mentioned either: Bernie Sanders, Hillary Clinton, Ted Cruz or Donald Trump. This particular stream is not parsed by which individual candidate was mentioned.

There is one very prominent node (Twitter user) on the upper portion of the twitter mention network for all of the candidates which may lead someone to think that information is disseminating from one source and is a hierarchical conversation. However, the rest of the graph tells a different story by revealing that there are multiple nodes (Twitter users) that have a high centrality within the network and that none of the candidates having a clustering coefficient higher than .0005. There are a lot of conversations occurring between different participants in the network which resembles a grassroots conversation. This graph also appears to also have a relatively high clustering coefficient due to the fact that a handful of nodes having a high-level of prominence and influence in comparison to other nodes. It was interesting to see that there were not any polarized crowds, which you might expect in the visualization of a network in which there

are two distinct political parties (republicans and democrats).

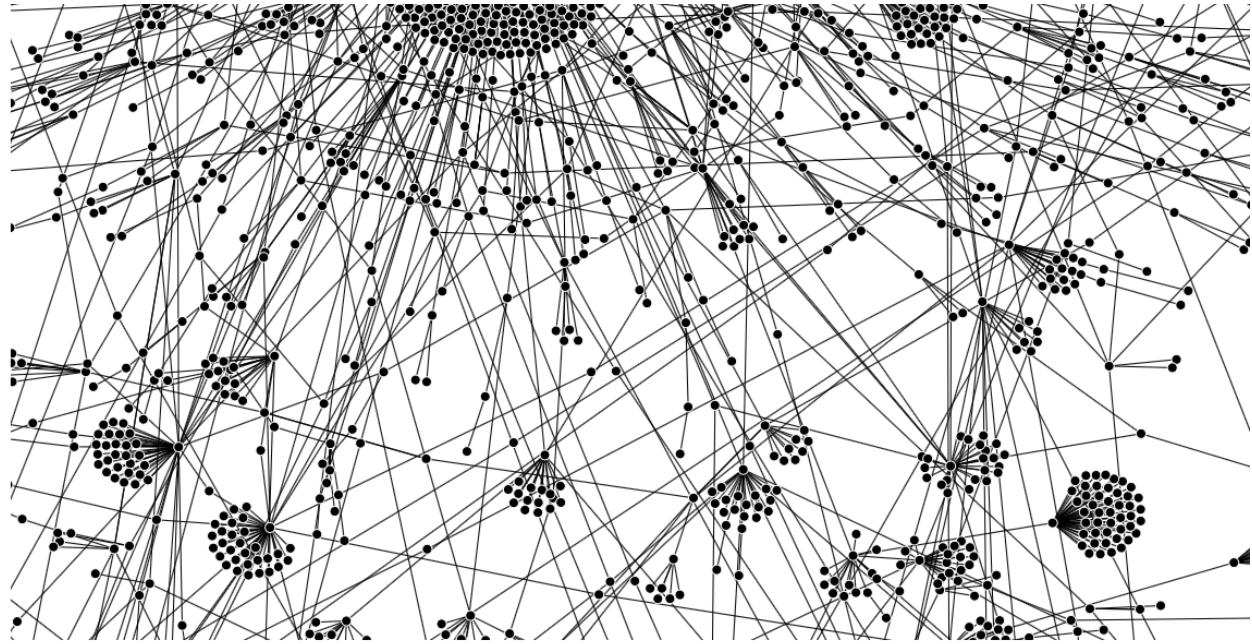


Figure 1 - Force-directed graph of the mention network for all four candidates (Bernie Sanders, Donald Trump, Hillary Clinton and Ted Cruz)

While there seems to be overlap between the Twitter mention networks for each candidate the visualizations of each candidate's twitter mention network reveal that the most influential users or nodes are different for each candidate. The most influential users are not the users that mention a lot of users but instead those that are most frequently mentioned and retweeted. The users that were identified as being the top mentioners of other users did not have a lot of traction in the network because their tweets were not retweeted or replied to whereas the most influential users are mentioned and included in conversation even when they are not tweeting anything. Additionally, there was not much continuity between which news sources seemed to be retweeted and mentioned most frequently for each candidate. The news sources did not serve as a source of centrality between the Twitter mention networks of the

candidates which reveals that News channels are not the most influential users in this Twitter mention network.

Additionally, for Donald Trump, Hillary Clinton and Bernie Sanders's mention networks the respective candidates were not the most retweeted/mentioned user but their official Twitter accounts were central to the network. In contrast, Ted Cruz had the smallest and least complex mention network (diameter = 2 and average path length = 1.014) and TedCruz was clearly the most prominently mentioned node. This suggests that Ted Cruz's twitter account, unlike the other candidate's, is the primary player in all of the twitter conversations regarding him. Whereas, for the other networks a lot of the conversation and buzz would occur even without the candidate's twitter account.

Force Layout Visualization of the Twitter Mention Network for Bernie Sanders

In Bernie Sander's network the users that were mainly mentioned were ViralBuzzNews , Bernie Sanders, Independent, DJ_Curfew, NewsBreaksLive, Time, ABC News and CNN. ViralBuzzNews appears on the left and it is interesting because while ViralBuzzNews has the most influential node in Bernie Sander's network it seems to be completely disconnected from the majority of the network. This finding reinforces the fact that this network cannot be defined as hierarchical because information is not disseminating throughout the network from one sole twitter user. Bernie Sanders is the second most influential Twitter user in his mention network. Additionally, there are a lot of smaller conversations which involve only a few participants that occurring about Bernie Sanders in the outskirts of the network visualization.

Sander's mention network had the widest diameter (5) in comparison to the other candidates. Trump's network diameter was 4, Clinton's was 3 and Cruz's was 2. It is interesting

that Sander's had the highest network diameter because Trump's network is much larger, however messages about Bernie Sanders seem to be traveling much further.

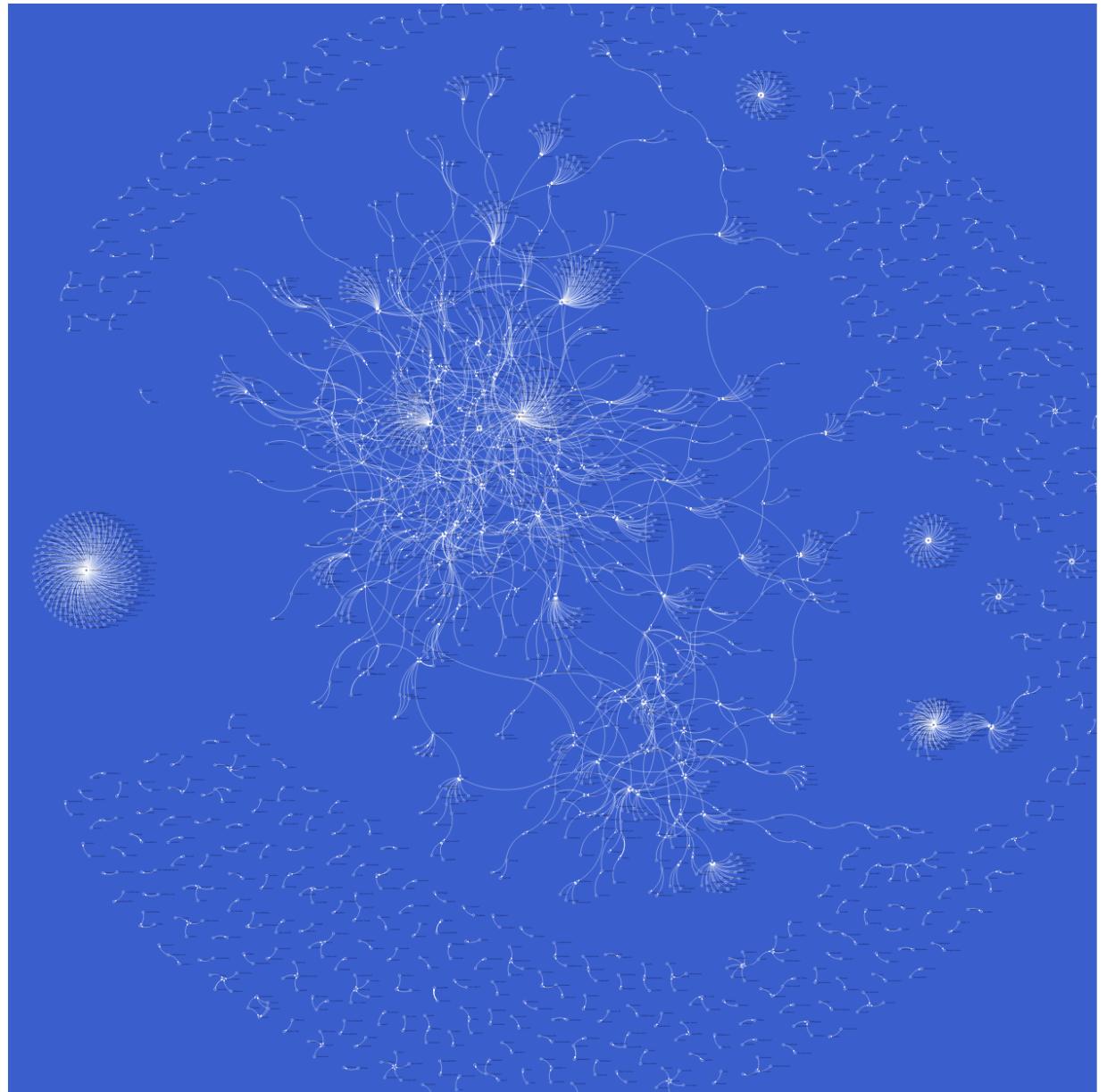


Figure 2 - Force-direct graph of the mention network for Bernie Sanders

Force Layout Visualization of the Twitter Mention Network for Ted Cruz

Ted Cruz's twitter network is the smallest out of the 4 candidates and especially pales in comparisons to Donald Trump who is currently his direct competitor in the Primary Presidential Election. The user that generated the most mentions of other user's in Cruz's network was girl_iowagirl20. The users that were mentioned the most in connection with Ted Cruz were tedcruz, voxdotcom, salon and deray. However, it should be noted that criteria used to identify the users that were mentioned the most within Cruz's network was much lower than the

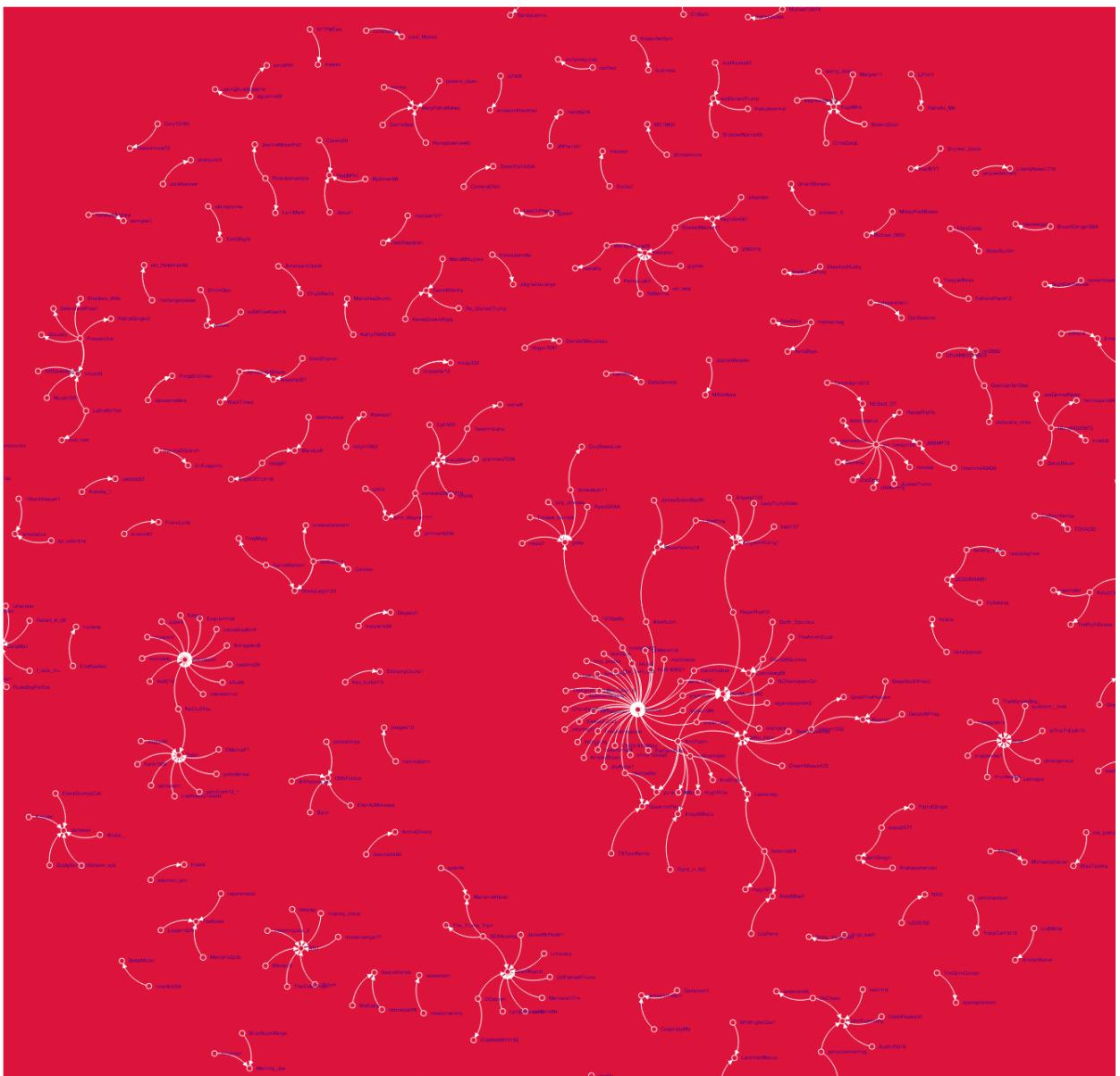


Figure 3 Force-direct graph of the mention network for Ted Cruz

threshold used for the other networks because the overall amount of interaction between any given user and the rest of the network was much smaller. Ted Cruz's network graph has a diameter that of 2 which is half the diameter of Donald Trump's. Additionally, Cruz's network also has a higher occurrence of lone pairs of nodes that are not part of triangles and leads to his network having a 0 clustering coefficient. This finding makes it appear as though there is very little hierarchy to the conversations that are occurring about Ted Cruz.

Force Layout Visualization of the Twitter Mention Network for Donald Trump

Donald Trump had the largest and most complex Twitter mention network therefore it was difficult to identify who generated the most tweets directed to other users within in his network. However, it was clear that some of the individuals that were most frequently mentioned were: creapills, realdonaldtrump, thedreamcloset, dory, levie, dahab17102015, cnner, bensfletcher, GOLDENGATEBLOND, OHNOSHETWIT, CARLROVE, NYCJIM, HIITAYLORBLAKE, SOURVOID, HAVEIGOTNEWS and . POWERFUL. This suggests that there are a lot of distinct conversations going on about Donald Trump simultaneously.

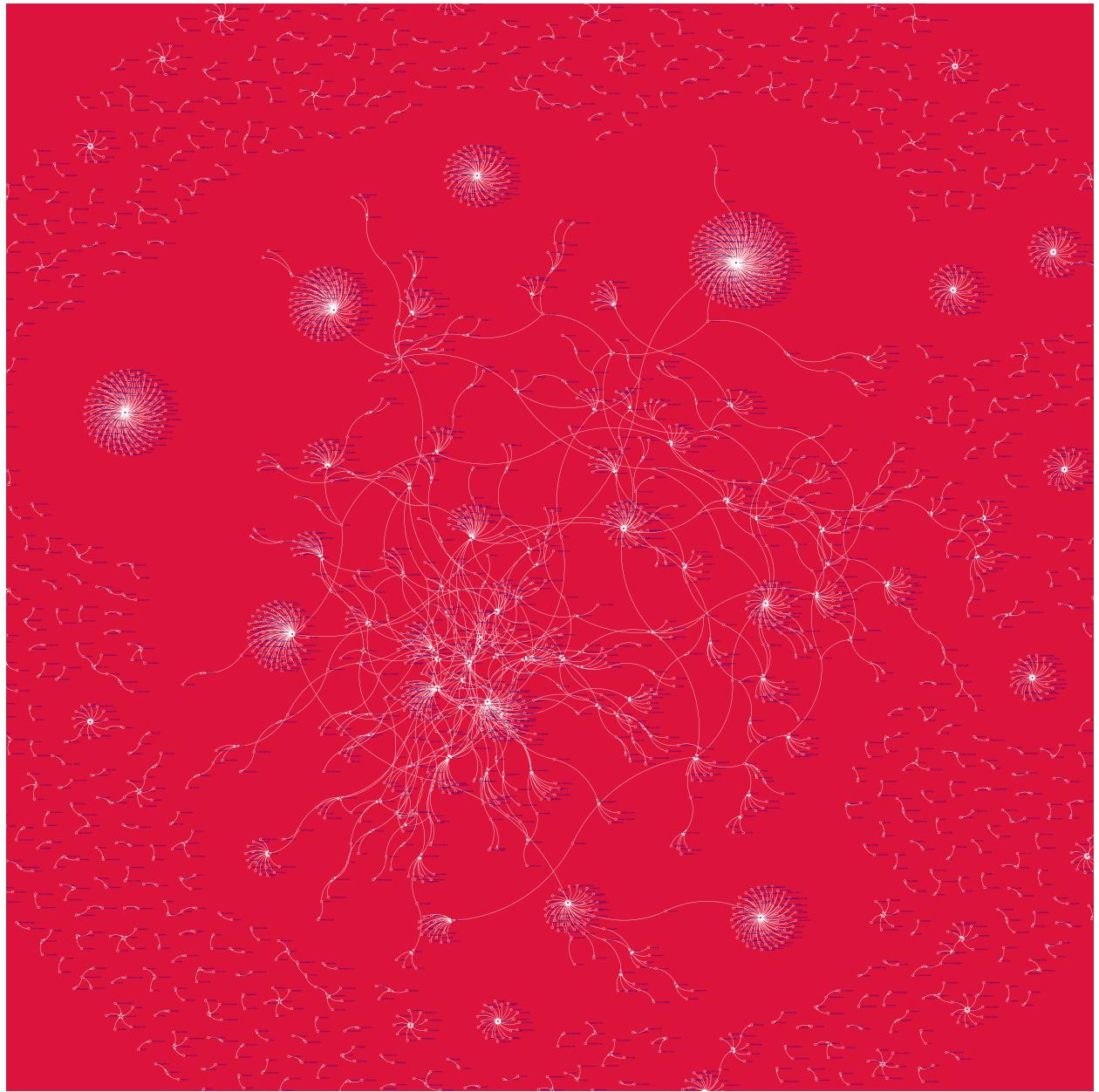


Figure 3 Force-direct graph of the mention network for Donald Trump

Force Layout Visualization of the Twitter Mention Network for Hillary Clinton

In Hillary Clinton's network some of the Twitter users that were mentioned the most were:

hillaryclinton, thehill, nadelparis, DJ_curfew, diplo and Britney spears. Britney Spears had a

much larger influence on Hillary Clinton's network than Hillary Clinton's twitter account.

Additionally, there was an interesting connection between two of the most mentioned users:

nadelparis and DJ_curfew which showed that their networks of tweets mentioning Hillary

Clinton were directly connected to each other independent of the rest of the graph. The person

who tweeted at the most people in Clinton's network was the user starrick1.

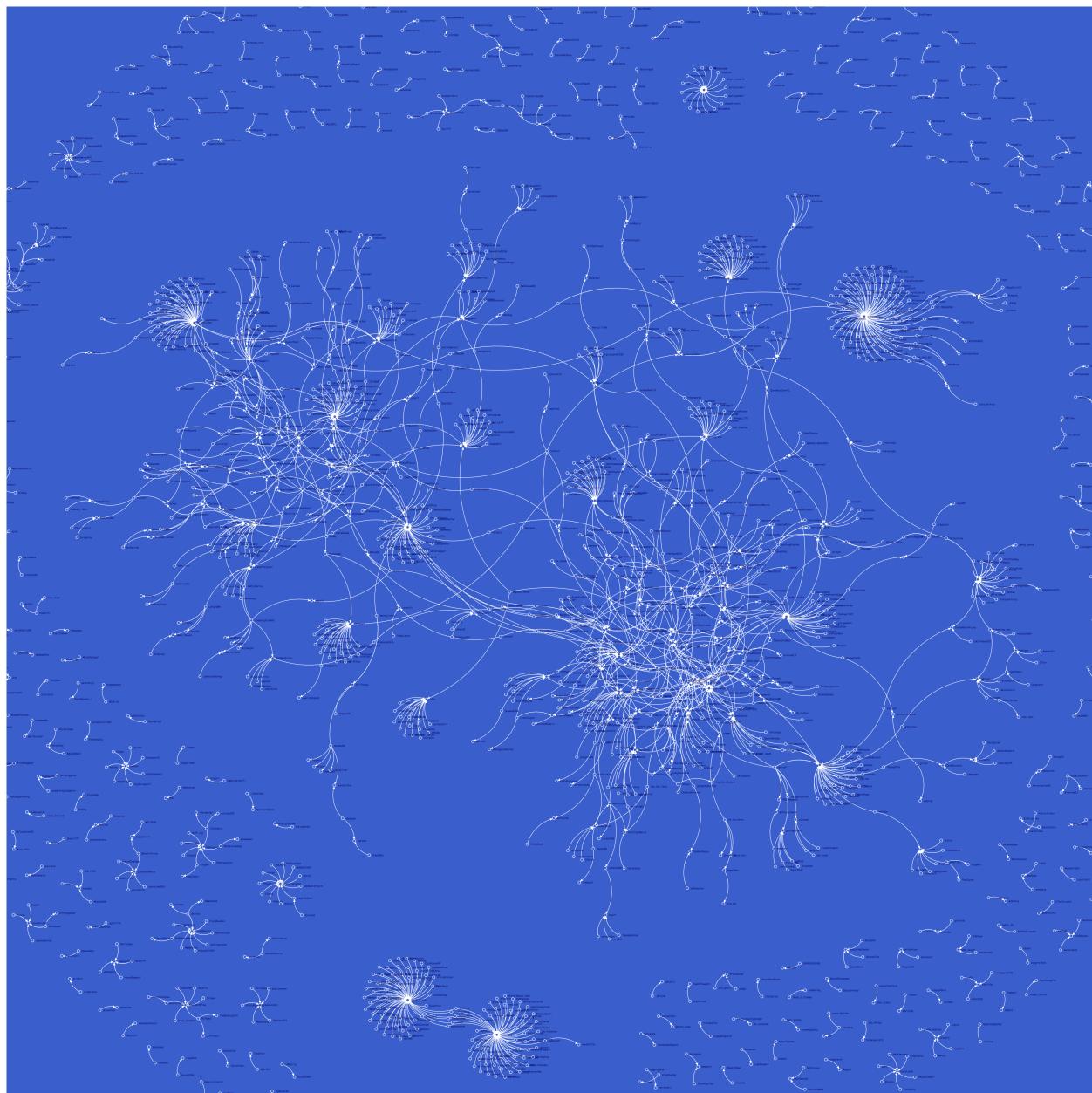


Figure 4 Force-direct graph of the mention network for Hillary Clinton

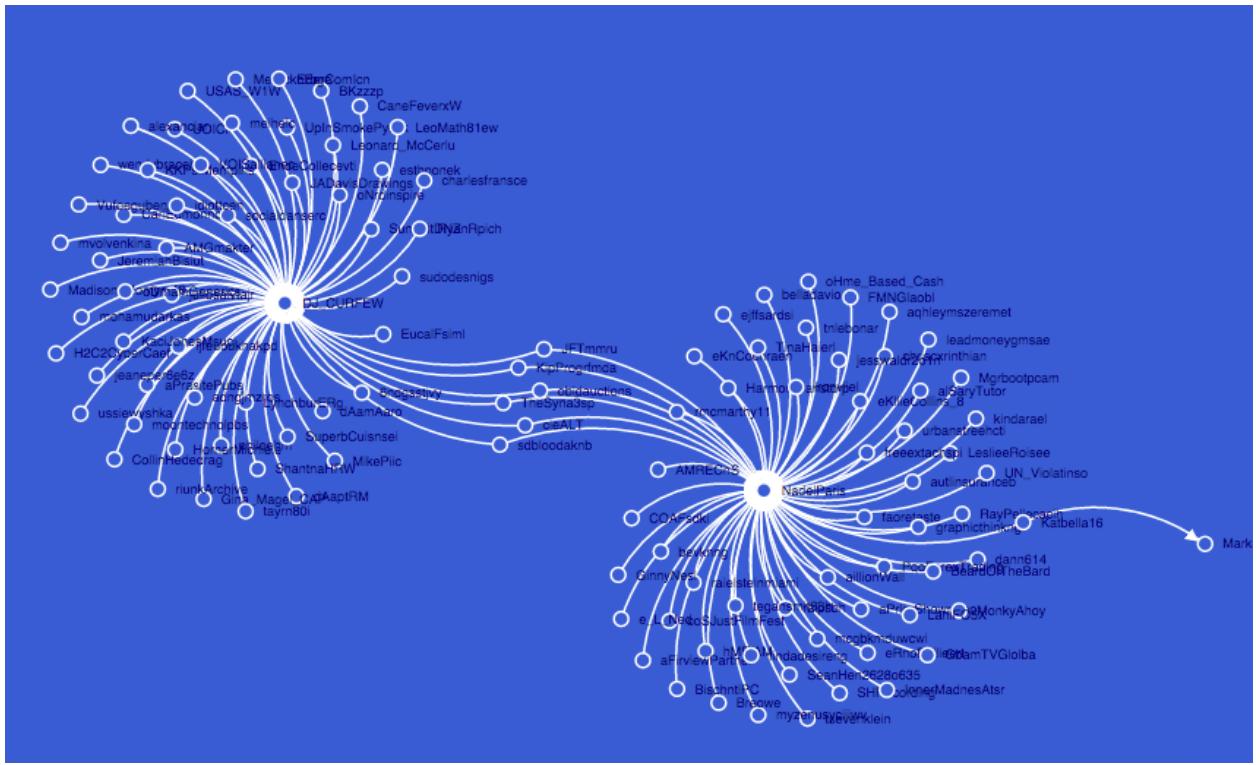
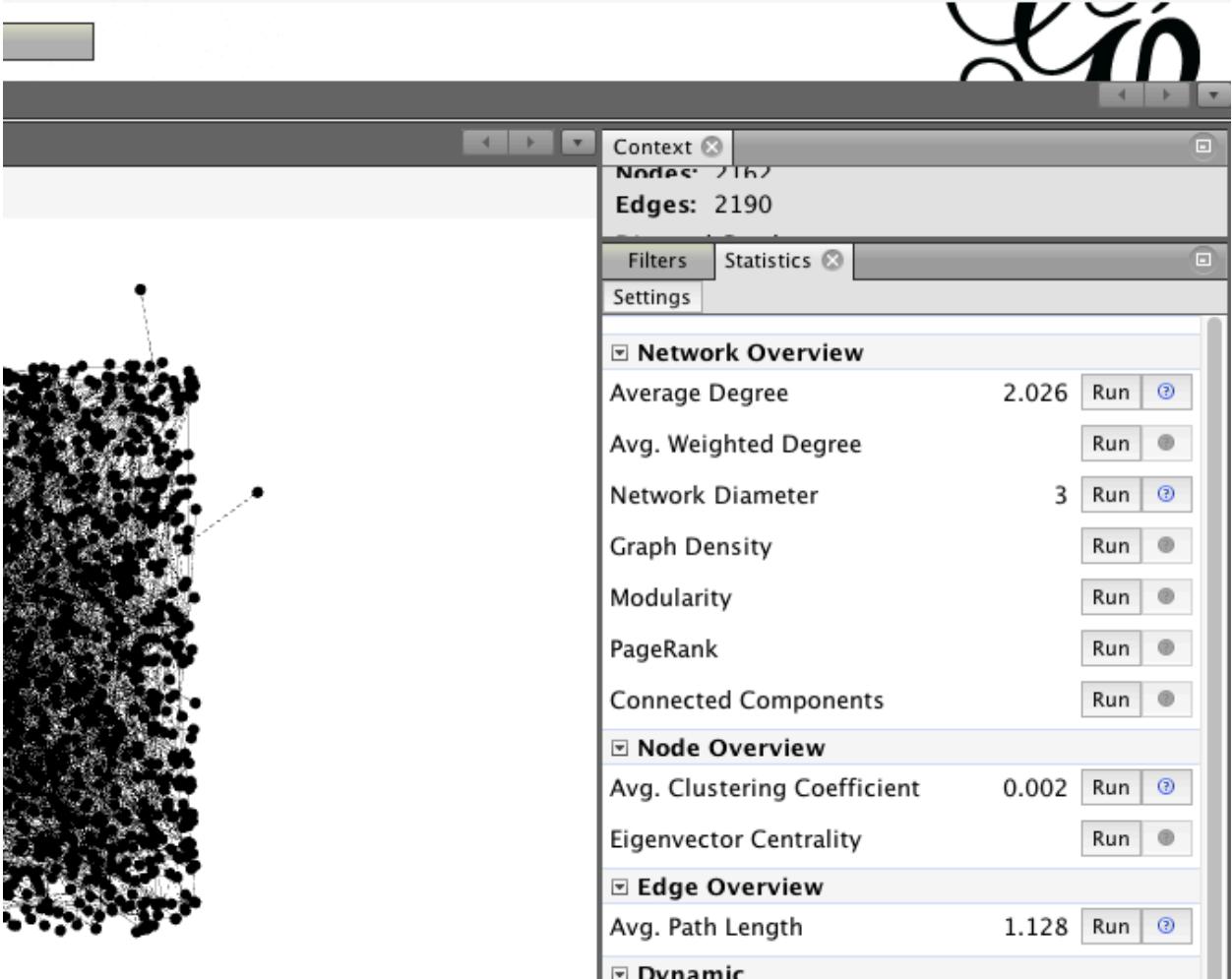
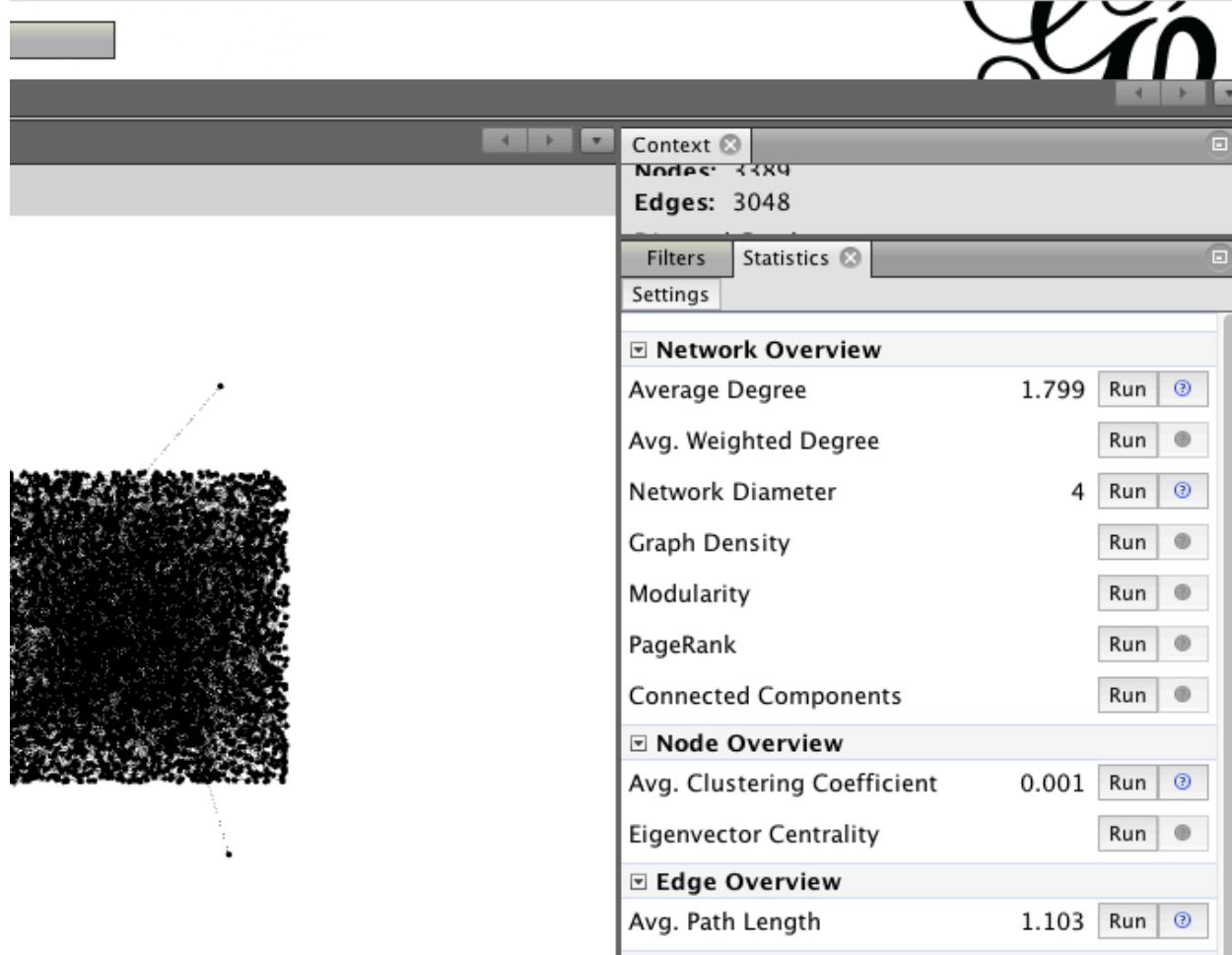


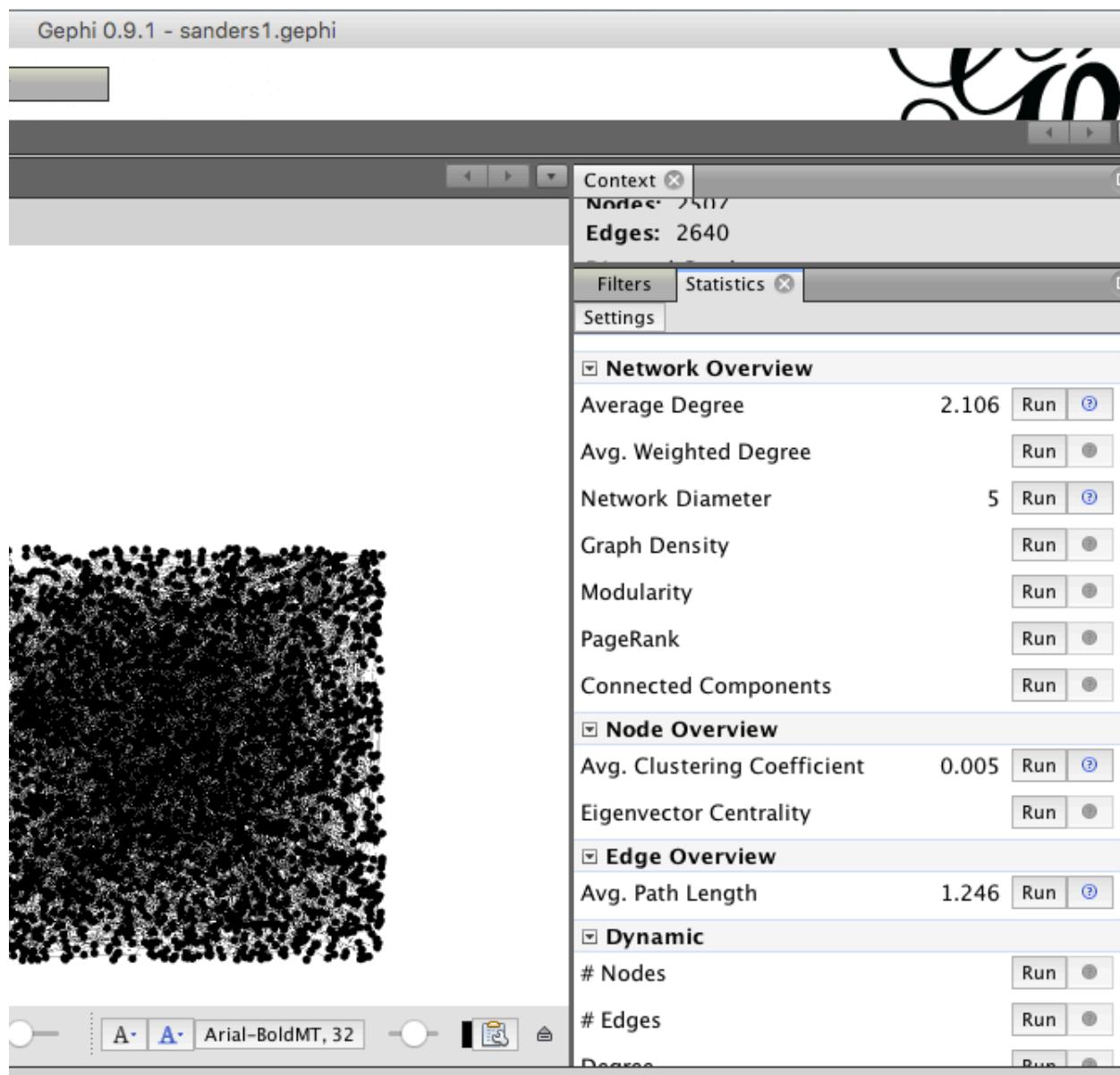
Figure 5 - Relationship between two influential users (Nadeparis and DJ_CURFEW) in Clinton's force-directed graph mention network

Gephi 0.9.1 - clinton1.gephi



Gephi 0.9.1 - trump_gephi1.gephi





Gephi 0.9.1 - cruz1.gephi

The screenshot shows the Gephi 0.9.1 software interface. On the left is a network visualization of a dense cluster of nodes connected by edges. The main window contains several panels:

- Context:** Shows "Nodes: 448" and "Edges: 360".
- Filters:** A tab in the top navigation bar.
- Statistics:** A tab in the top navigation bar.
- Settings:** A tab in the top navigation bar.
- Network Overview:** A section with the following metrics:
 - Average Degree: 1.607 (Run button)
 - Avg. Weighted Degree (Run button)
 - Network Diameter: 2 (Run button)
 - Graph Density (Run button)
 - Modularity (Run button)
 - PageRank (Run button)
 - Connected Components (Run button)
- Node Overview:** A section with the following metrics:
 - Avg. Clustering Coefficient: 0 (Run button)
 - Eigenvector Centrality (Run button)
- Edge Overview:** A section with the following metric:
 - Avg. Path Length: 1.014 (Run button)
- Dynamic:** A section with the following metrics:
 - # Nodes (Run button)
 - # Edges (Run button)
 - Degree (Run button)

At the bottom, there are font size and style buttons (A+, A-), a font selection dropdown (Arial-BoldMT, 32), and other interface elements.