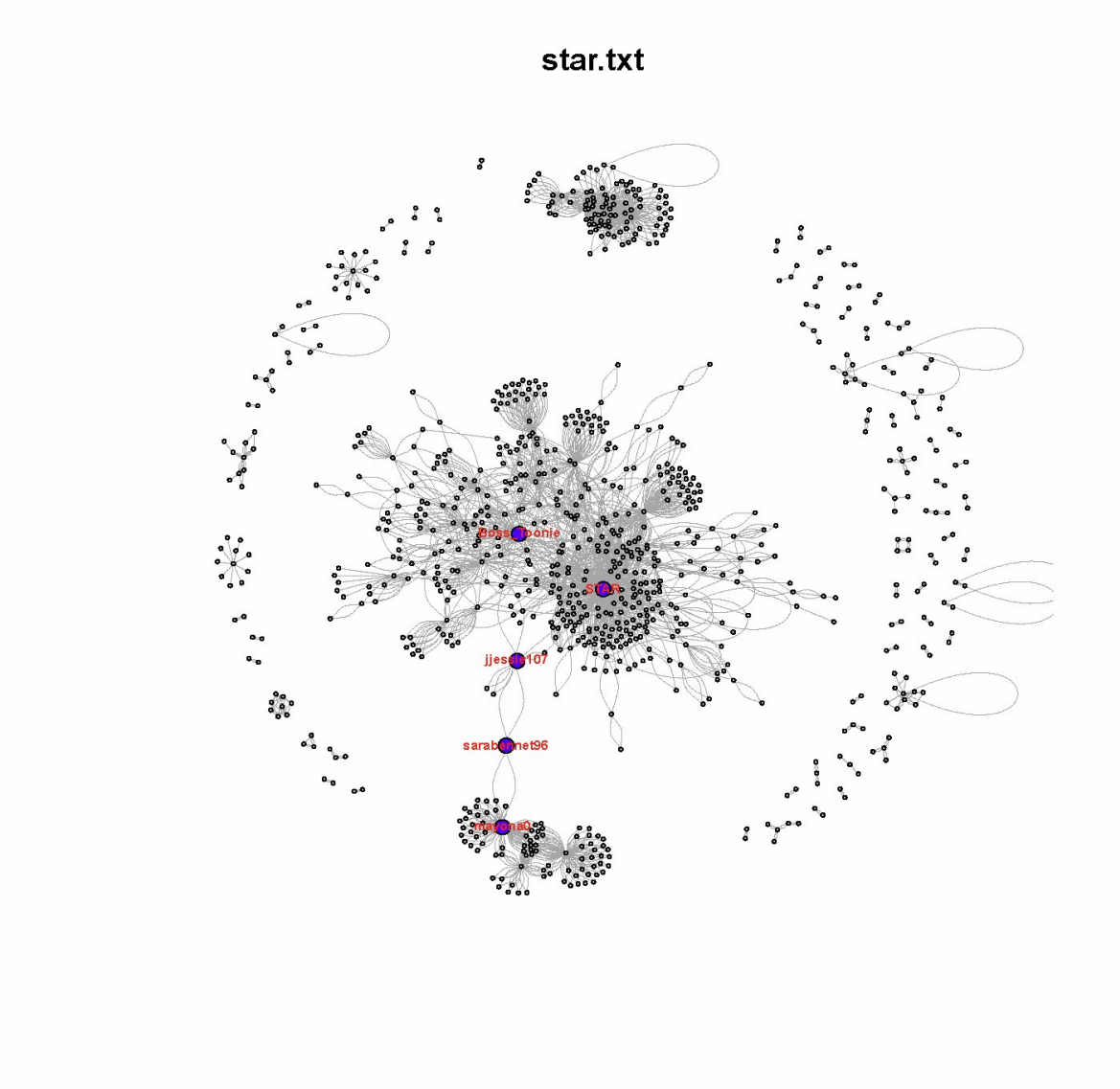
Professor Yang Feng from the University of Virginia's College published an article titled "Are you connected..." (2015) which analyzed the flow of information on Twitter regarding the Starbucks #RaceTogether campaign. In this article, the author created a network of users connected via retweets/mentions and determined the accounts with the top 10 betweenness centrality values. Of those 10 users, he identified 5 main types of Tweeters:

1. **Conversation Starter**: the account that started the online discussion (i.e. @Starbucks)
2. **Active Engager**: someone who retweets/mentions frequently
3. **Influencer**: someone who gets retweeted/mentioned frequently
4. **Information Bridge**: someone who connects Active Engagers and Influencers through retweets
5. **Network Builder**: someone with high betweenness centrality but doesn't retweet often nor gets retweeted often.

I was curious to see if these 5 types of tweeters could be found in many retweet/mention networks, or just in this instance. So I collected tweets with the hashtag #RogueOne, which was the hashtag @starwars was using to promote their new movie. I collected tweets for about an hour during the evening. I also collected tweets for the hashtag #STAR (for the premier of the TV show Star) and tweets for the hashtag #Empire (for the season finale of the TV show Empire).

To collect the tweets I used Twitter's Streaming API. From those tweets I made a network similar to the one Professor Feng used. The nodes are twitter handles (i.e. @STAR). I put an edge from node A to node B if A mentions or retweets B. Before I looked at the Tweets, I predicted that the tweets/tweeters probably would not follow the "5 main types of Tweeters" framework because I thought that retweets would be too chaotic. It also seemed more likely to me that users would be both Active Engagers and Influencers (meaning that they retweet often and get retweeted often) instead of just one of the two.

The first network I analyzed was the #STAR network. Running the tweets through my R script generated this graph:



Drawing 1: Retweet Graph for #STAR

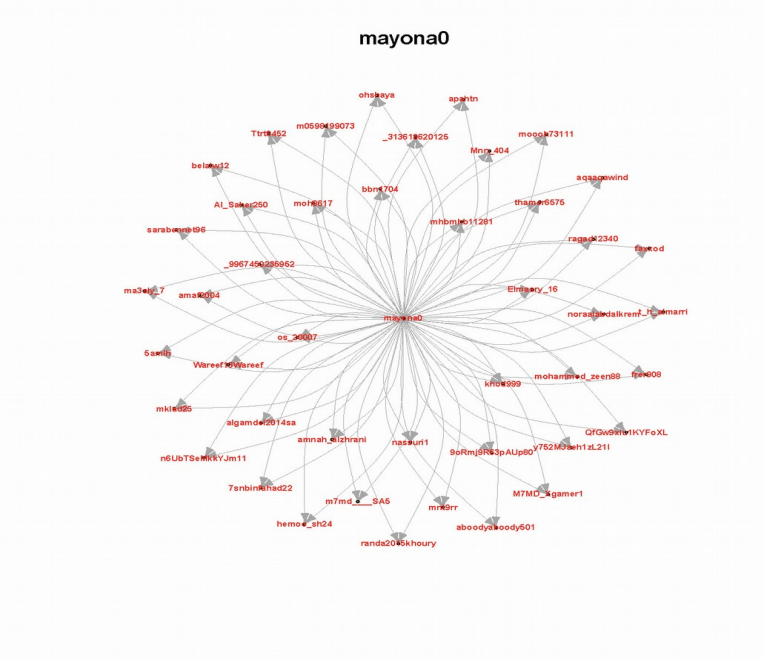
I will attach the full resolution image of all the graphs with this report. It's impossible to read the node names even if the graph is as large as the page itself. Analyzing this graph gave the following statistics:

```
| "Top betweenness values"
      STAR   jjessie107      mayona0 sarabennet96  Boss_Toonie
 123131.84     44486.13     43936.72     42484.00     34910.02
| "Top betweenness values (in largest connected subgraph)"
      STAR   jjessie107      mayona0 sarabennet96  Boss_Toonie
 123131.84     44486.13     43936.72     42484.00     34910.02
| "# times they were retweeted (out-degree)"
      STAR   jjessie107      mayona0 sarabennet96  Boss_Toonie
       407            11           88            0            0
| "# times they retweeted (in-degree)"
      STAR   jjessie107      mayona0 sarabennet96  Boss_Toonie
        29             0            0            4           61
```

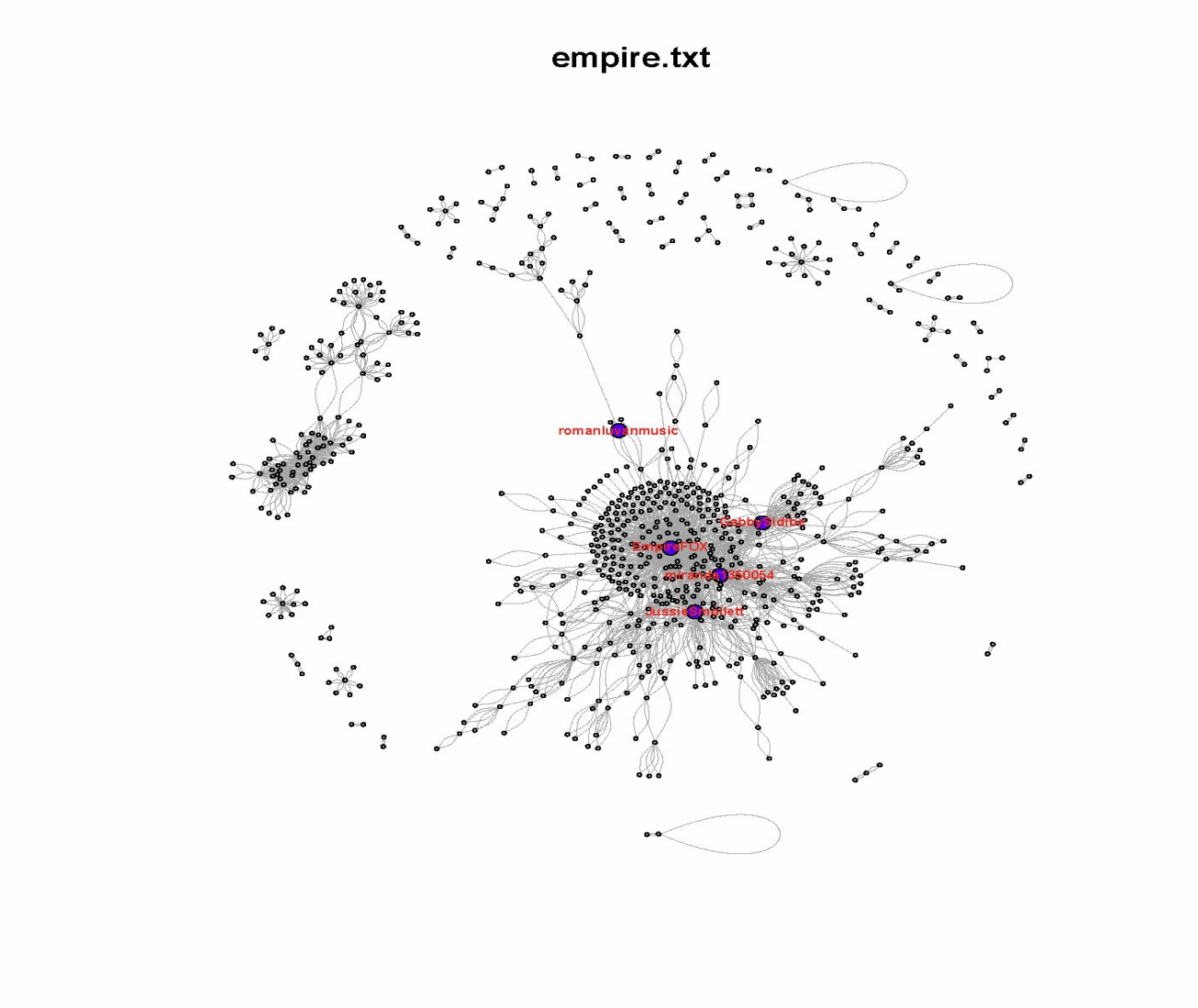
Unsurprisingly, the @STAR account had the highest betweenness centrality of 123131.84. I had to make the graph undirected to generate these numbers. If the node has all arrows pointing towards or all arrows pointing away, that node will have a betweenness of zero, since nothing can "flow through" that node. However, we want to include such a node in our analysis so I set "directed" to "FALSE". The rest of the users are normal (non-celebrity) people.

@STAR is by definition the Conversation Starter for the network because they're the account for the TV show. @jjessie107 and @mayona0 are acting like Influencers because they were retweeted/mentioned very often but didn't retweet/mention other users at all. Everyone seems to be listening to what they have to say. On th other hand, @Boss_Toonie seems to be the Active Engager because he/she retweeted others 66 times but was never retweeted by others 0 times. I interpret @Boss_Tonnie's function in this network as someone who is pushing information forward in the network, without generating any new information. Finally, @sarabennet96 seems be acting as a Network Builder and Information Bridge. By definition she is the former because she is in the top 5 betweenness centralities even though not many retweets passed through her. She is the latter because she connects the Influencers (@jjessie107 and @mayona0) with the Active Engager (@Boss_Tonnie), which you can clearly see in the graph.

Drawing 2: Retweet graph for @mayona0



The next network I analyzed was the #Empire network. Here's the graph and statistics:



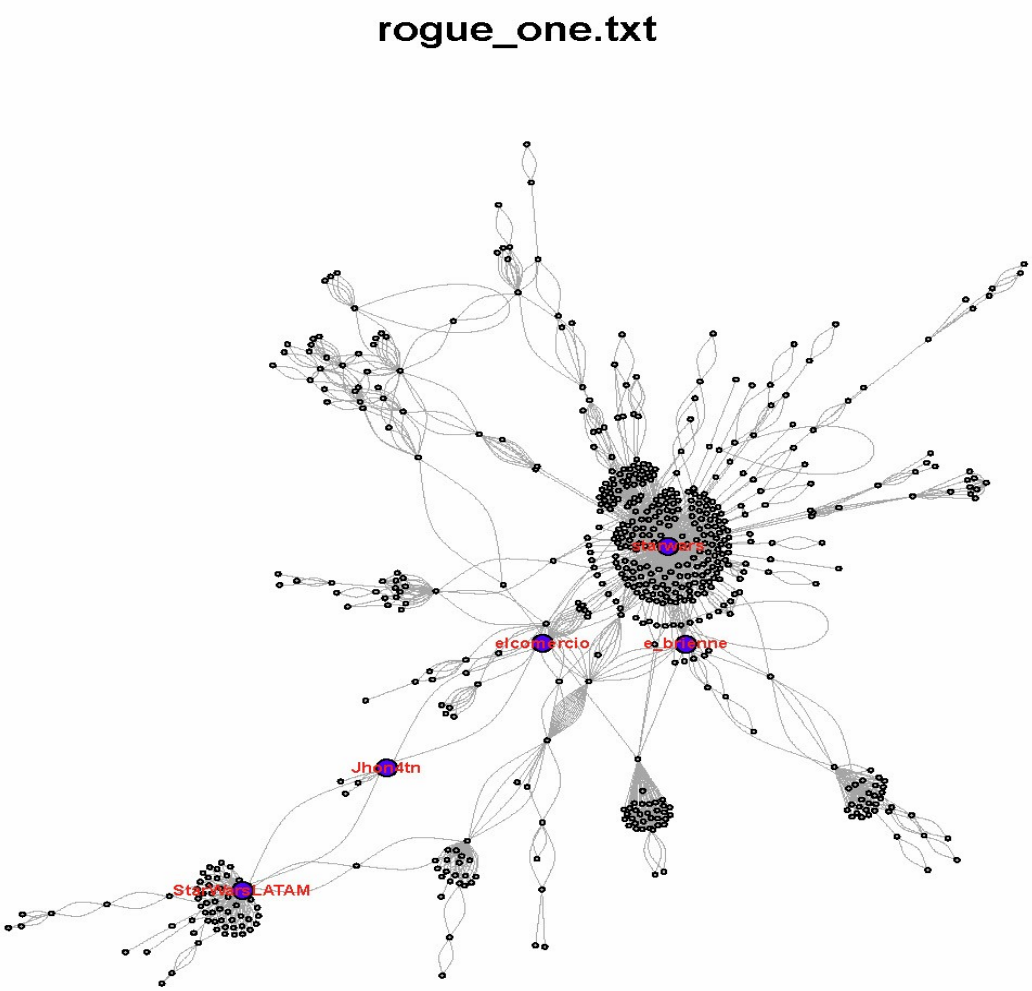
Drawing 3: Retweet Graph for #Empire

```
[1] "Top betweenness values"
      EmpireFOX     GabbySidibe  miranda1350054  JussieSmollett romanluvanmusic
     101480.662       12352.806       10610.057        9207.891        8702.000
[1] "Top betweenness values (in largest connected subgraph)"
      EmpireFOX     GabbySidibe  miranda1350054  JussieSmollett romanluvanmusic
     101480.662       12352.806       10610.057        9207.891        8702.000
[1] "# times they were retweeted (out-degree)"
      EmpireFOX     GabbySidibe  miranda1350054  JussieSmollett romanluvanmusic
            776              71               0             129               0
[1] "# times they retweeted (in-degree)"
      EmpireFOX     GabbySidibe  miranda1350054  JussieSmollett romanluvanmusic
             47               1              67               0               2
```

In this graph, @EmpireFox is obviously the Conversation Starter. The Influencers are @GabbySidibe and @JussieSmollett because they were retweeted frequently. The Active Engager is @miranda1350054 because that person retweeted others very often. Finally, @romanluvanmusic was the Network Builder because he was retweeted 0 times and retweeted others only 2 times, yet had the fifth largest betweenness centrality in the graph. There didn't seem to by any Information Bridge type user.

The last network I analyzed was the #ForceOne StarWars network. Let's look at that graph and the statistics:

Drawing 4: #RogueOne Graph



```
"Top betweenness values"
 starwars StarWarsLATAM    elcomercio       Jhon4tn     e_brienne
154884.92      32702.14      31395.14      31231.14      25178.29
"Top betweenness values (in largest connected subgraph)"
 starwars StarWarsLATAM    elcomercio       Jhon4tn     e_brienne
154884.92      32702.14      31395.14      31231.14      25178.29
"# times they were retweeted (out-degree)"
 starwars StarWarsLATAM    elcomercio       Jhon4tn     e_brienne
      515           104            14             0             0
"# times they retweeted (in-degree)"
 starwars StarWarsLATAM    elcomercio       Jhon4tn     e_brienne
        0             0             2             8            20
```

The #RogueOne graph was so big that I've only included the largest connected subgraph, and not the entire thing. This graph didn't fit the 5 user framework quite as well as the other two examples. @StarWars and @StarWarsLATAM are the Conversation Starters. @elcomercio seems to be an influencer, but the number of times retweeted isn't as large as in other examples. @jhon4tn and @e_Brienne seem to be Active Engagers. While the data in this example isn't as good as in other examples, I still find the results meaningful. The fact that we were able to locate users that exclusively retweet others and (almost) exclusively get retweeted simply by examining users with high centralities is surprising to me.

In conclusion, I think these 3 examples seem to support the 5 user framework. In the first two examples, nearly all the high betweenness centrality Twitter users seem to fall very strongly into one of the 5 categories. I definitely did not expect this result, as I predicted that Twitter users would be too chaotic to categorize in such a way. It would be interesting to see if this result holds on other hashtags as well.