ADVANCED LAB 5

SIDDHESH TIWARI

UNIVERISTY OF ILLINOIS AT CHICAGO

UIN: 657796780

Link Prediction: Inferring the NASDAQ 100 Network

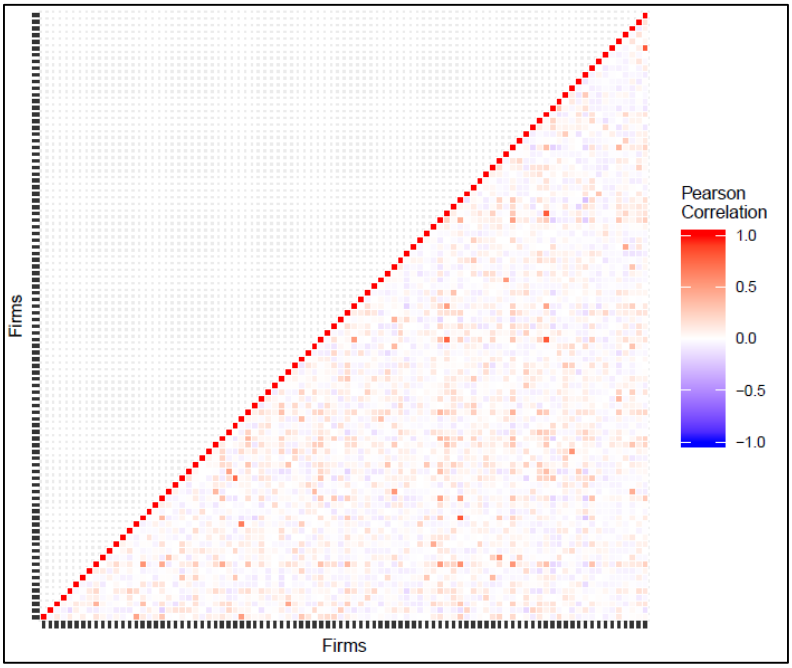


Fig 1 Lower Triangle Correlation Matrix of Firms

There are very few number of firms that show a strong correlation (positive or negative) in their twitter activities and among these most of the firms have strong positive correlation. These firms which exhibit strong positive correlation can be expected to form some kind of network. Let's look more closely at the positively correlated firms.

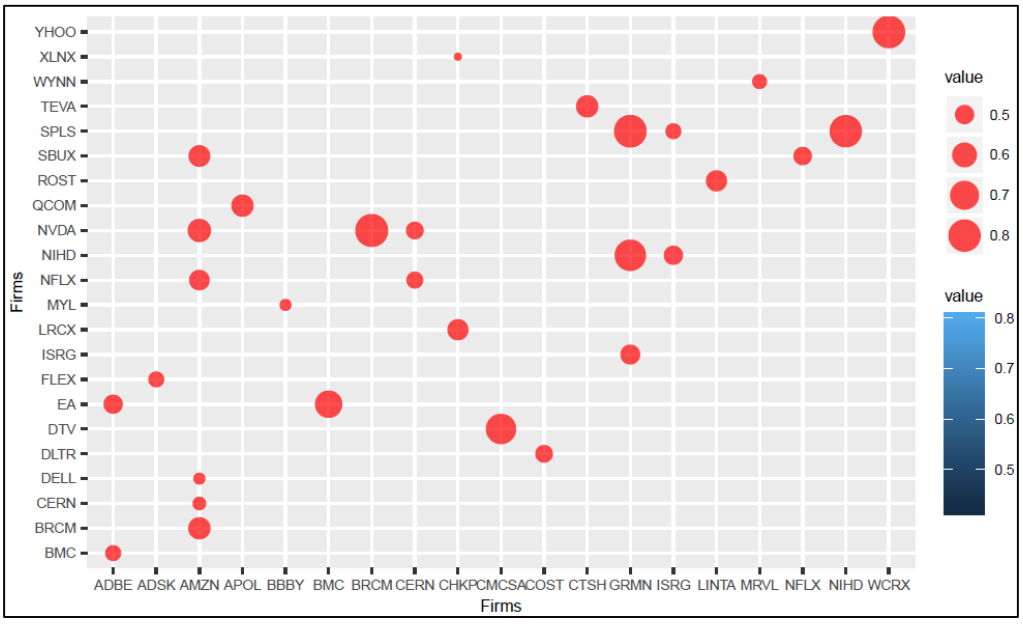


Fig 2 Positively Correlated Firms with Correlation between 1 and 0.4

After looking at the above plot we can anticipate that the correlation patterns can lead to generation of some sort of network between firms which can help us to analyze what are the type of firms which are correlated and what does there edges signify.

As we can see from Fig 2 Amazon's twitter activity is highly correlated with that of Nvidia, Starbucks, DELL and Netflix, we can expect that there is some kind of underlying network hidden in this correlation patter. We can infer similar kind of behavior from other correlated firms. Let's check the predicted network based on significant correlation and see if it can support our inference.

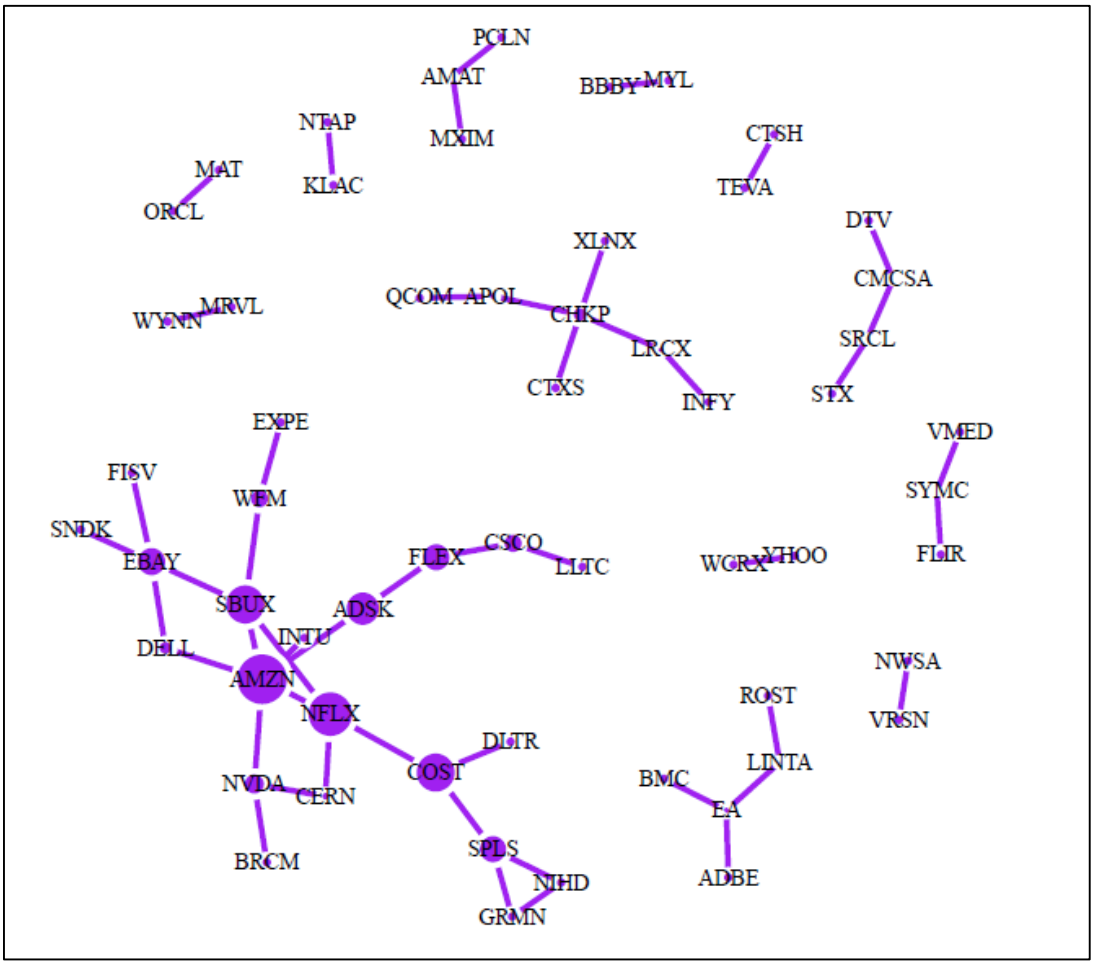


Fig 3 Network of NASDAQ Firms at 0.05 significance level

The prediction of links based on significant partial correlation yield 49 edges between nodes. Examination of the above network we can confirm the inference we were concluding from Fig 2 is represented in network. We can clearly see a network between Amazon, Netflix, Nvidia and other firms based on correlation from Fig 2.

Let us first define what will be considered as Competitive and Cooperative relationship. Competitive relationship will be one in which firms with same business domain are connected and Cooperative relationship will be one in which business domain are different. From this point of view if we look at the biggest sub-network we can see that main players in this network are Ebay, Amazon, Netflix, Costco & Staples, all of these firms belong to some sort of retail or content provider business. The existence of this sub-network can tell us that there is some kind of competitive relationship between these firms though we cannot be sure of it but this insight can lead to further investigation of their relationship. In the same sub-network we see that there are firms who are basically product manufacturer like Dell, Sandisk, Cern, Nvidia, Garmin and others are Financial, Payment, Discount, SCM & Network service provider like Fiserv Inc, Intuit, Dollar Tree, Flextronics, CISCO. There is no direct connection between these firms and they are

mostly connected to main retail players which indicates some kind of cooperative relationship for eg. Dell, Sandisk, Garmin, Cern might be cooperating to sell their products on retailers outlet or portals and firms like Intuit, Flextronics, Fiserv, are providing supporting functions to the business of retailer so that it can run smoothly.

Now let's look at the second biggest sub-network, this network consist of firms like Qualcomm, Citrix, Lam Research, Xilinx & Infosys these companies are either Software or Hardware service providers or both. So we can sense their might be some competitive relationship between them which will only be clear after deep investigation. There also other small bi-node network like one between Cognizant(CTSH) & Teva Pharma Industries(TEVA) which might be due some cooperative relationship in which Cognizant might have undertaken a Huge project from Teva. Another cooperative relationship can be one between Applied Material(AMAT) and Maxim Integrated(MXIM) where Applied material supplies equipment, services and software to enable the manufacture of semiconductor chips and Maxim Integrated designs, manufactures, and sells analog and mixed-signal integrated circuits.

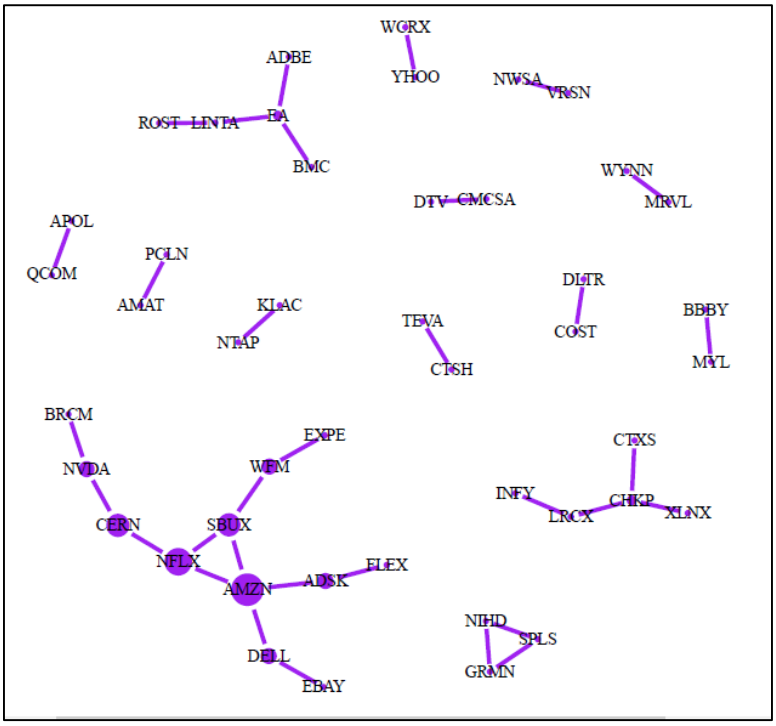
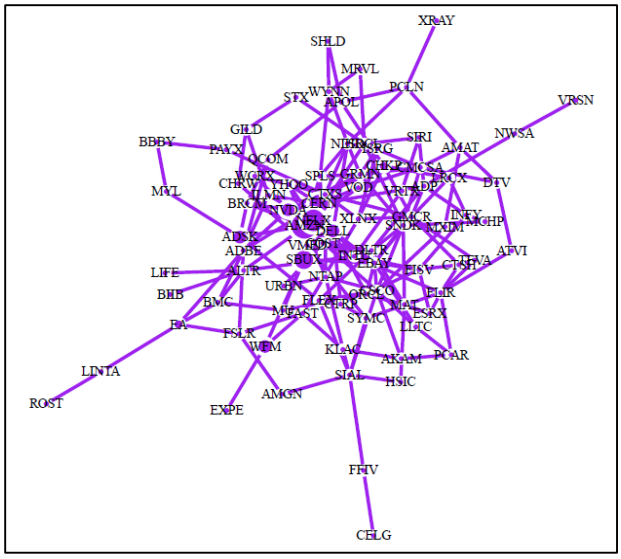


Fig 4 Network of NASDAQ Firms at 0.01 significance level

The prediction of links based on significant partial correlation at 0.01 significance level yield 33 edges between nodes. As we can see from above network graph that there has been reduction number of edges. If we compare it with Fig 3 we will find that largest connected component from previous network containing Amazon, Ebay, Staples and Costco has been decomposed here with Staples and Costco having their own small network. Similarly the second largest sub-network from previous graph has also been decomposed into smaller networks.



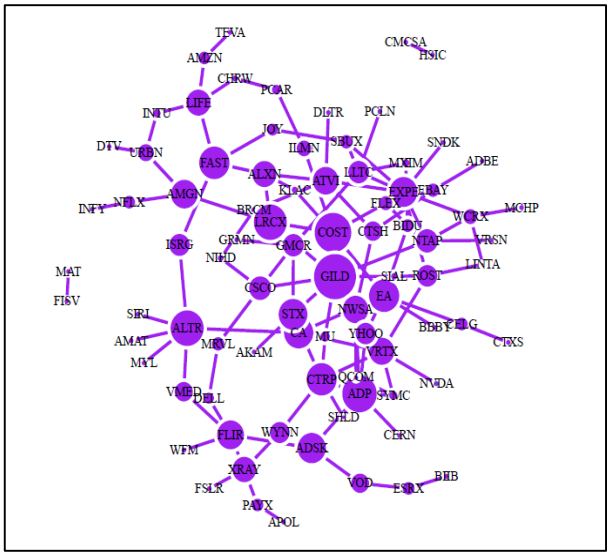


Fig 5 Network of NASDAQ Firms Using HUGE

Fig 6 Network of NASDAQ Firms Using Correlation

Looking at the network graph generated from HUGE with 109 edges and from using overall correlation values with 114 edges we can see the graphs are not very insightful as compared to partial-correlation values. Therefore, our best bet should be on partial-correlation to get good insights from network.