5000 Years of Notable Lives: Measuring Influence Across Cultures
Professor Etienne Wasmer - Instructor Palaash Bhargava
Final Project: The Network of Influence in the Tech Industry
Alyssa Yu, Ahmed Meshref, Muhammad Anas Jawed, Yuliya Frolova
December 12, 2018

## The Network of Influence in the Tech Industry

Overview

According to Bloomberg, people are connected to each other only if they are currently on the same company boards. However, there is a plethora of ways people may know each other; through familial ties, friendships or from having worked together in the past. We wanted to account for these types of connections, and realized that we could do it by using data from Wikipedia biographies. The idea was to build an interactive network where people would be connected to each other only if their Wikipedia biographies suggested they knew each other in some capacity.

In order to achieve our goal, we used multiple tools. We wrote scripts that used the Python library BeautifulSoup to parse the hundreds of Wikipedia pages we were working with, and used NetworkX to actually build the network node by node and edge by edge. In order to properly visualize the network and carry out analysis, we exported the NetworkX graph in the '.gexf' format, and edited it using the program Gephi.

Methodology

We started the process by identifying 'Key Executives' of the top 20 companies from the Bloomberg data, and then noted down the Wikipedia links for the people who had them through manual researching. We found that out of a total of 278 key executives, only 76 had Wikipedia pages (see 'Key People.csv').

Stage 0: We began building the network by writing a Python script (s01.py) that read through the CSV file containing names of the key people and links to their Wikipedia pages and made a simple graph based on them. In order to make sure that we were using BeautifulSoup and NetworkX correctly, we limited the goal of this step to the following: for each key person who has a Wikipedia page, go to their page, extract the name Wikipedia
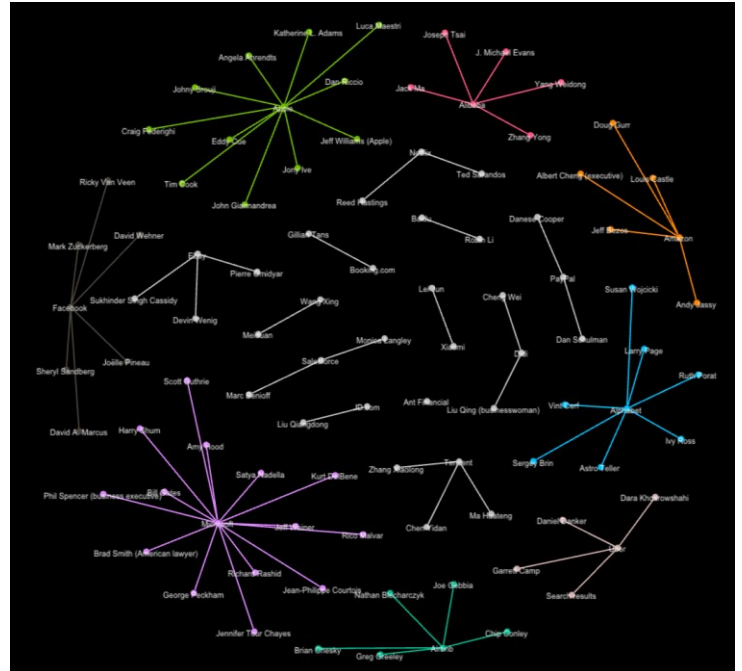
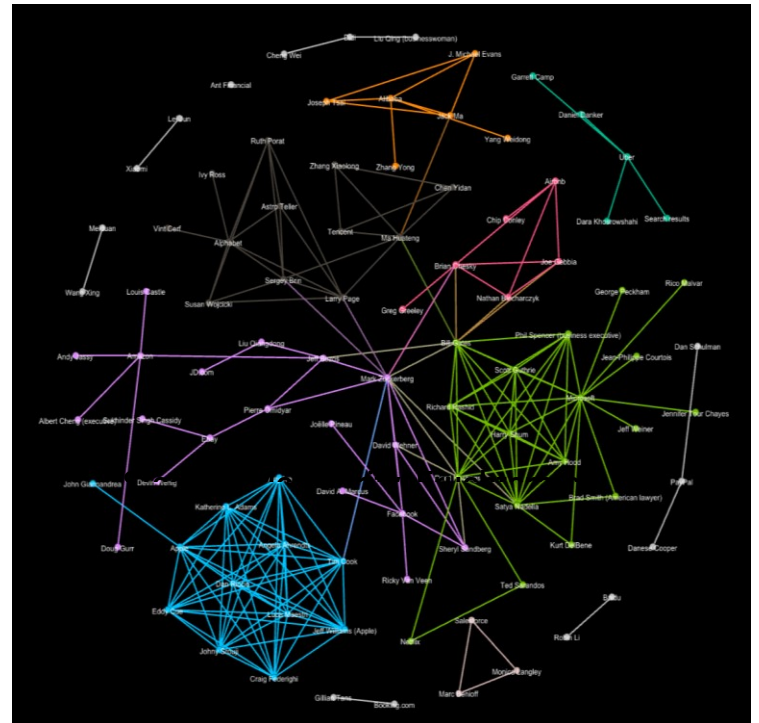

Stage 0: Modularity, Fruchterman Reingold

uses for them, make a node with this name and connect it to the node that represents the company they belong to. At this stage, we had a total of 96 nodes (76 people, 20 companies) and 76 edges.

Stage 1: We now wanted to see if our key people were related to each other directly. For this, we added a few lines to the script (s01.py) which checked, when adding a node, if the 'bodyContent' of the page represented by the node contained a link to any of the other key nodes. If this was the case, we added an edge which directly connected the two nodes.

By doing so, the number of nodes of course remained the same, but the number of edges jumped to 196. As is visible, most of these new links were within the company

clusters (as people working for the same company know each other), but there were also some new connections between people of different companies.

    We ran our analysis first by using the measure of modularity, which tests the strength of the connections within clusters in contrast to nodes in other clusters. In the force atlas graph below, the clusters with high modularity are color coded accordingly. These clusters resemble company groupings with which individuals are affiliated (e.g. light blue: Apple, light green: Microsoft, orange: Alibaba). This unsurprisingly displays the intra-connectedness within individuals who work in the same company compared to those who work with different companies.

**Stage 1: Modularity, Fruchterman Reingold**

Stage 1: Modularity, Force Atlas Layout
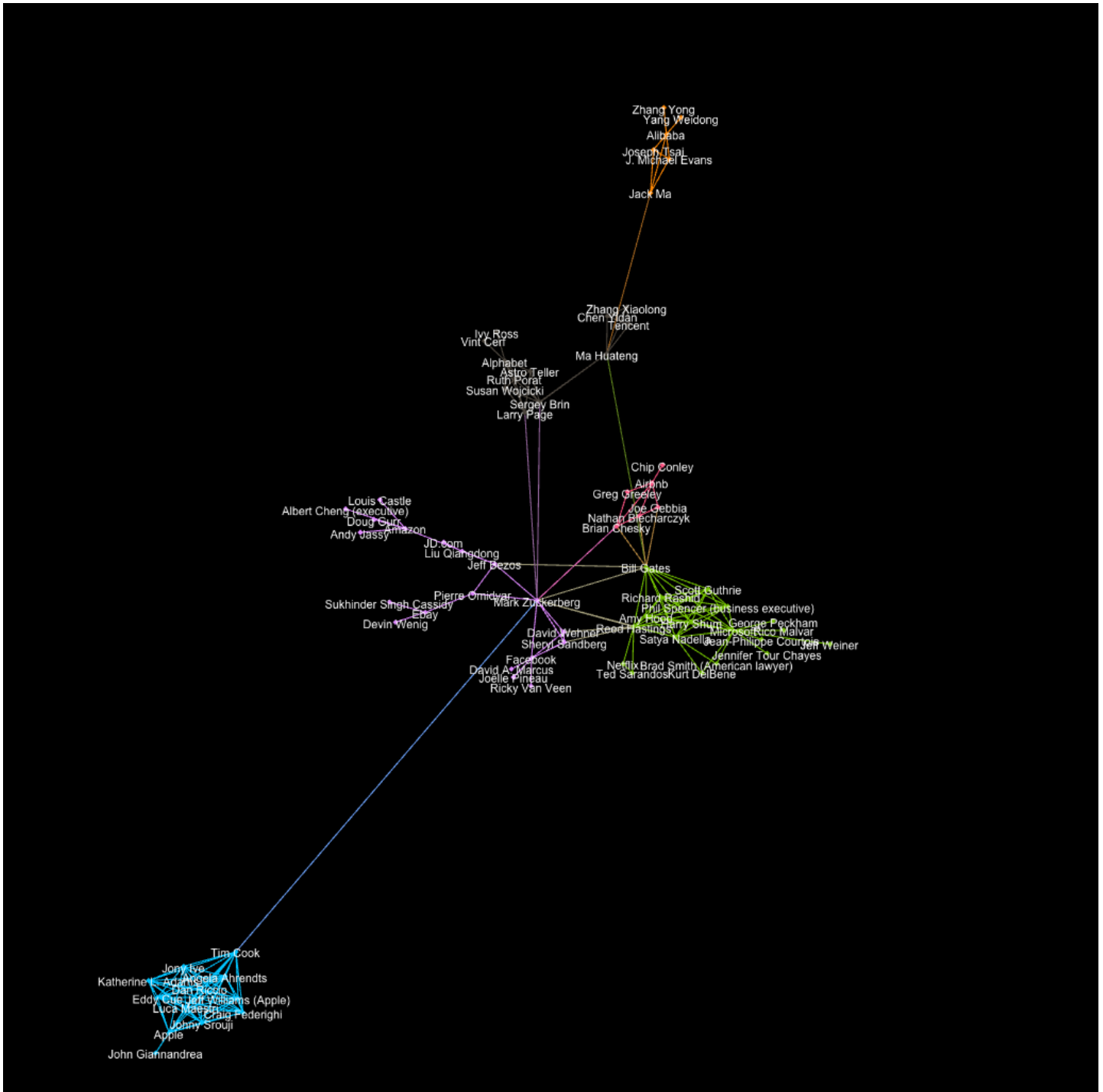
However, it is also important to note that despite the strength of these company

intra-connections, the different company clusters display some connections with each other too. In order to examine this further on a more individual scale, we ran our analysis by using the degree of betweenness centrality of individuals. This refers to the centrality of individuals, through which the most number of shortest paths pass through to connect two other people. In this context, we liken these central individuals to intermediaries that bridge gaps between their company and other companies. From the graph below, the darker green lines, which represent higher degrees of betweenness centrality, trace back to Facebook CEO Mark Zuckerburg, Apple CEO Tim Cook, Microsoft founder Bill Gates, Tencent CEO Ma Huateng, to name a few. This is particularly evident in the case of Apple, which is displayed in the bottom left hand side, where it remains detached from the rest of the clusters, save for the connection between Apple CEO Tim Cook and Facebook CEO Mark Zuckerburg. Chinese companies Tencent and Alibaba mainly lie in the peripheries towards the top. JD.com on the middle left hand side is seen connected to Jeff Bezos and members of Amazon perhaps due to similarities in e-commerce platforms provided. Therefore, though these companies are mostly intra-connected with people within their companies (reflected in clusters comprised of white lines), their company leaders serve as intermediaries that connect their companies with other large tech companies.

Stage 1: Betweenness Centrality, Force Atlas Layout
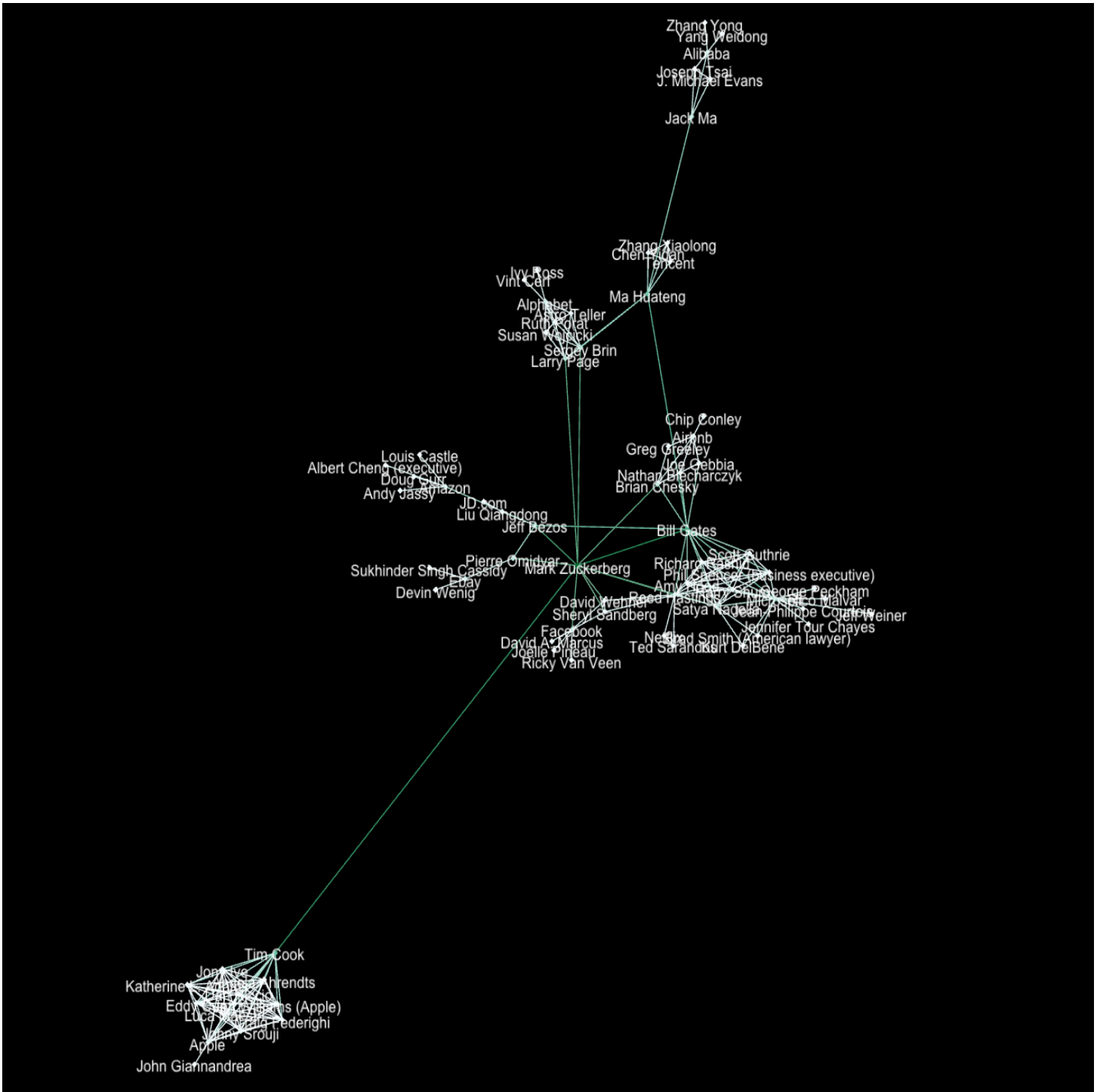
Stage 2: We now wanted to see if any of our original 'Key People' were connected to each other via other notable technology professionals that we had not accounted for. Using a script kindly provided to us by Palaash Bhargava and Professor Etienne, we extracted the names and Wikipedia links for 771 people categorized under 'American Technology Chief Executives' (found here, and saved in file the700.csv). We
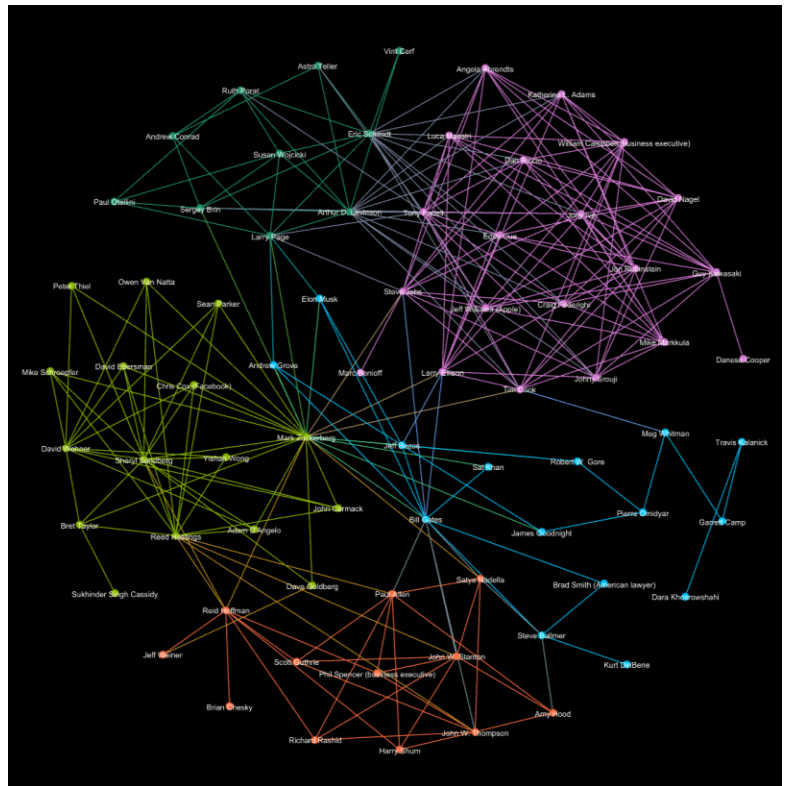
Stage 2: Modularity, Fruchterman Reingold Layout

then modified our code to check, for each key person, whether there were any people from the700.csv database which connected them to other key people.

In the process, we discovered 35 new people/nodes who served as 'intermediaries' in connecting our Key People. The total number of nodes now became 131, and the number of edges rose to 432. Some notable intermediaries we found are: Steve Jobs, Elon Musk, Sal Khan and Paul Allen. The full list can be found in the document 'New People.docx', and the code is stored in the file 's2.py'.

Analysis

Table 5.1. Most Connected Executives (Source: Wikipedia built by Wiki Nodes)

Name	Company	Position
Mark Zuckerberg	Facebook	Chief Executive Officer
Reed Hastings	Netflix	Chief Executive Officer
Bill Gates	Microsoft Corporation	Principal founder
Tim Cook	Apple Inc.	Chief Executive Officer
Luca Maestri	Apple Inc.	Senior Vice president and Chief Financial Officer
Jeff Williams	Apple Inc.	Chief operating officer under CEO Tim Cook
Katherine L. Adams	Apple Inc.	General Counsel and Senior Vice President of Legal and Global Security
Angela Ahrendts	Apple Inc.	Senior Vice President
Eddy Cue	Apple Inc.	Senior Vice President of Internet Software and Services

|		
Craig Federighi	Apple Inc.	Senior Vice President of Software Engineering

Table 5.2. Bloomberg Ranking of the Most Connected Executives (Source: Bloomberg built by Wiki Nodes)

| Name	Company	Position	Bloomberg Ranking
Mark Zuckerberg	Facebook	Chief Executive Officer	8 Relationships
Reed Hastings	Netflix	Chief Executive Officer	40 Relationships
Bill Gates	Microsoft Corporation	Principal founder	45 Relationships
Tim Cook	Apple Inc.	Chief Executive Officer	22 Relationships
Luca Maestri	Apple Inc.	Senior Vice president and Chief Financial Officer	3 Relationships
Jeff Williams	Apple Inc.	Chief operating officer under CEO Tim Cook	No Relationships
Katherine L. Adams	Apple Inc.	General Counsel and Senior Vice President of	No Relationships

		Legal and Global Security	
Angela Ahrendts	Apple Inc.	Senior Vice President	13 Relationships
Eddy Cue	Apple Inc.	Senior Vice President of Internet Software and Services	19 Relationships
Craig Federighi	Apple Inc.	Senior Vice President of Software Engineering	No Relationships

Comparing the above to Bloomberg, we see that the key executives who were listed by Bloomberg as having the most number of relationships do not directly match the key executives the networks we created by scraping Wikipedia notes as being the most well connected.

The addition of former Apple CEO Steve Jobs as an intermediary created connection to Google Co-Founder Larry Page, Salesforce CEO Marc Benioff, and
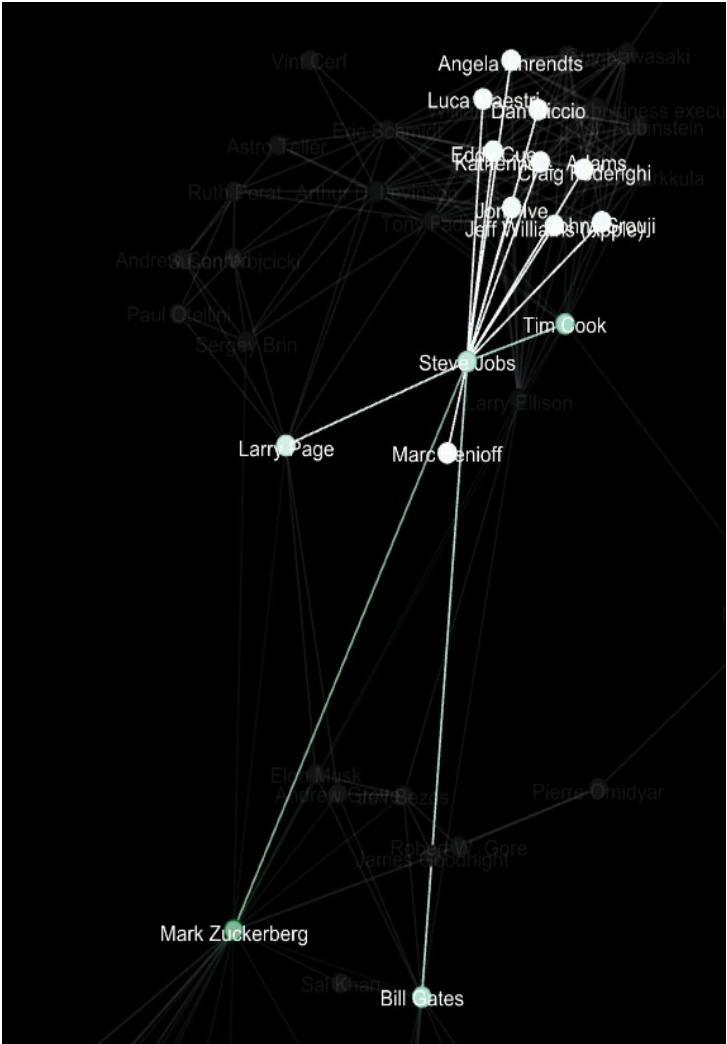
Microsoft Founder Bill Gates that were not present previously with current CEO Tim Cook.
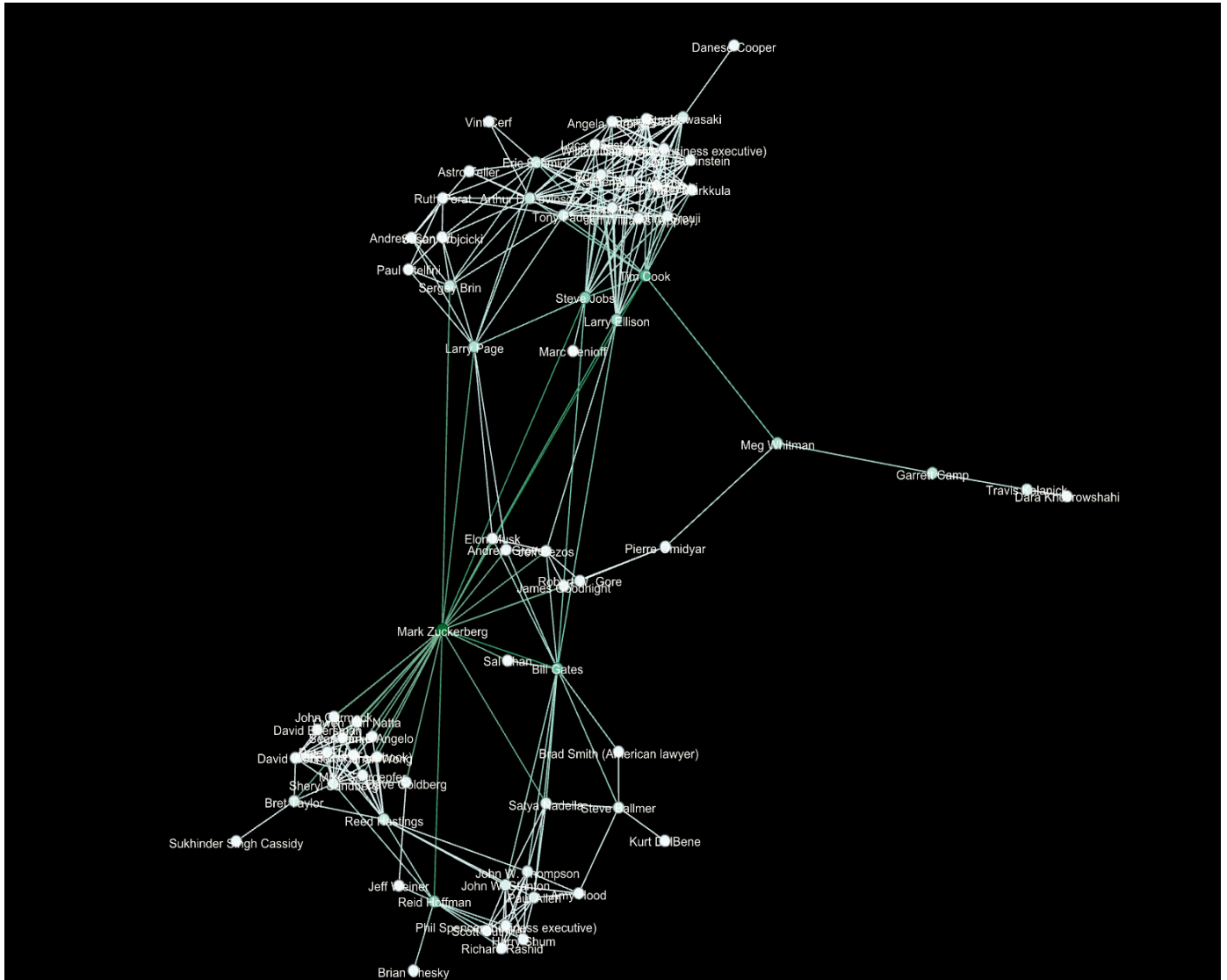
We found that despite the presumably little number of intercompany connections between Chinese and American tech companies, the Chinese companies themselves were hardly connected with one another.

We ran a similar analysis conducted previously by extracting the names and

Wikipedia links for 560 people categorized under 'Chinese Business Executives' (found [here](), and saved in file 'Chinese Executives.csv'). We then checked each key person to see if they would serve as intermediary intercompany connections. Although this list reinforced intracompany connections, no external intercompany connections were formed. The lack of these connections may reinforce our hypothesis regarding the centralized nature of Chinese tech companies, where power and notability (as exhibited in the few pages available) are concentrated in a few key players. However, this could partly also be attributed to a limitation of the English language edition of Wikipedia that had less Wikipedia pages for Chinese key executives.

Conclusions

Performing network analysis based on several information sources, such as Bloomberg and Wikipedia, allowed us to build different sets of nodes and edges. As the literature and research on networks would suggest it is difficult to draw the exact bounds of connectivity level, intra-companies vs inter-companies. Key executives and board members have various ways in which they might be related as well: on personal (family/friends) level, alumni (alma mater) relations, conferences/talks, co ownership etc. However, the general trend within companies and its key administratives can be tracked. American companies tend to have more edge links within the industry, while Chinese companies stay very clustered on the single company level. From the comparative tables of top-5 executives by each source, we can see that through developed and highly edited Wikipedia pages, people who are highly active on social

media platforms as well as are active participants of the Internet/tech sector innovations, are highly connected through both primary job and social life networks. While Bloomberg presents more information through company's insider perspective, providing detailed connectivity of the key executives based on their company management work and investment decisions.

Table 6.1. and 6.2. Comparative tables of Wiki Nodes vs Bloomberg Relationships Ranks

BRs = Bloomberg Relationships

Rank	Name	Company	Wiki Nodes	BRs
1	Mark Zuckerberg	Facebook	31	8
2	Reed Hastings	Netflix	26	40
3	Bill Gates	Microsoft Corporation	24	45
4	Tim Cook	Apple Inc.	22	22
5	Luca Maestri	Apple Inc.	20	3

Rank	Name	Company	BRs	Wiki Nodes
1	J. Michael Evans	Alibaba Group	141	3
2	Daniel H. Schulman	PayPal	135	1
3	Vinton G. Cerf	Google	130	3
4	Yun Ma	Alibaba Group	128	4
5	Jeffrey Weiner	LinkedIn	122	3