“What influences a friendship network on facebook?”

By Charles Du

Motivation

The Facebook, a social networking site focused on the college community, was created by Mark Zuckerberg while he was an undergraduate in Harvard. Since its release in 2004, it has taken college campuses by storm and has become one of the most checked webpages by college student. Due to its prevalence (8 out of 10 college students check the facebook daily) and high quality of information (only people with an .edu address can have an account), the facebook network is a great candidate to study.

My goal for this study is to map out my friendship network and try to find the influencing factors that lead to friendship network clusters. Some of the factors that I considered are:

   Hypothesis: Students who study in the same department will share many school-specific experiences including classes, homework sessions, clubs that will lead to friendship networks in facebook.

2. Year – 1st, 2nd, 3rd, 4th, graduate
   Hypothesis: Students who are in the same year will share many year-specific experiences including living in dorms, orientation, and rush that will lead to friendship networks in the facebook.

3. Origin – International student or American Citizen.
   Hypothesis: Students who are international will share many international student-specific experiences including international orientation, international student clubs that will lead to friendship networks in the facebook.

By mapping out my social network and running a clustering analysis using the girvannewmantool, I will be able to see which of the factors stated above are important in forming a friendship network. I will also be vigilantly watching other factors that may be important for cluster formation.

Related Work

A related study was done by Lada Ademic on the Club Nexus network (http://www.firstmonday.dk/issues/issue8_6/adamic/).
In her study of “A Social Network Caught in a Web,” she performed an analysis on how people’s interests influenced their perceptions of themselves (surveys asking “what kind
of person do you think you are”) and the friendship networks they formed. Lada writes: “…in general, activities or interests that are shared by a smaller subset of people showed stronger association ratios than very generic activities or interests that could be enjoyed by many….” The study that I will be conducting will be focused on how factors like department of study, year in school, and origin (citizenship or not a citizenship) will influence friendship networks.

Data Collection
In order to map out the social network, I used a feature of the facebook that allowed me to see connections among all my first degree of friends.

Figure 1 Facebook's Mutual Friends listings shows 1st degree connections
Since Facebook did not allow web crawlers to be used, I manually entered nodes and edges information into an Excel sheet and exported the spreadsheet into a CSV file that I guess could use.

The resulting network contained 155 nodes and 603 edges.

![Figure 2 Building a nodes and edges file](image)

![Figure 3 Mapped Network with Kamada – Kawai arrangement](image)
Results of Analysis

According to the graphs plotted in GUESS and rearranged by the Kamada-Kawai algorithm, a person’s department of study seems to be the most influencing factor when it comes to making friends. Figure 4 below shows the friendship network colored-coded with different departments of study.

As you can see with figure 4 below, there seem to be two distinct groups with a majority of yellow nodes (engineers) on the right side and a majority of red nodes (LS & A) on the left side. When you look at the links that each node, you’ll notice that the left connected node (from the department of engineering) links to many other engineers while vice versa happens with the node from the department of LS&A. This makes sense because in any given department, you will spend a lot of time together with other people in the same department taking the same classes, doing homework assignments, doing projects.

![Network showing different schools of study](image)

Figure 4 Network with showing different schools of study
We can confirm the grouping by areas of study using the girvannewmantool algorithm to discover clusters. As depicted in figure 5 below, two clusters emerge with a clear separation by school of study.

Figure 5 Cluster analysis confirmed that area of study is a strong influencer in forming clusters

Further use of the girvannewmantool algorithm revealed new clusters and exposed new factors that influenced friendship networks. The new cluster below shows that many people where connected to their friends from the same high school. It makes sense in this because I went to a high school in Michigan and a significant number of mutual friends that I knew in high school came to the University of Michigan.

Figure 6 Common high school formed a new cluster
Another significant cluster that was discovered contained a network of people that lived in the same dorm (East Quad). This makes sense because dorms are very densely populated that facilitate many friendships to occur.

![Cluster by Dorm](image)

**Figure 7** A friendship network around people who lived in the same dorm

More clusters exposed the significance of shared activities on friendship networks. Two clusters shown below grouped people that went to the same church (same spirituality) and people who played soccer together (same sport). These clusters support Lada’s theory that activities and interests shared by people tend to form friendships.

![Cluster by Other Affiliation](image)

**Figure 8** Clusters showing people who share similar activities
Additional clusters were discovered to form around a common event called Leadershape, an intense 1-week leadership workshop that created many friends. An interesting subcluster of only aerospace engineers was exposed in the engineering cluster.

![More Clusters](image)

**Figure 9** Leadership Camp and Aerospace Engineering formed more clusters

Surprisingly, year in school and origin did not strongly influence the formation of clusters. The rectangles with letters below represented non U.S. citizens and the different colors represented different years. The reason could be that people in different years and different origins could still share similar classes and activities which where the strongest clustering effects occurred.

![Year and Int Status were not significant](image)

**Figure 10** Students in different years and different origins did not group together
An analysis was also done looking at the degree distribution for each person. No significant correlations were found between the school of study and the number of degrees. This means that your selected major does not play a major part in determining the size of your friendships network.

![Figure 11 Degree Distribution](image)

An analysis was also done looking at the betweenness distribution for each person. No significant correlations were found between the school of study and the betweenness.

![Figure 11 Betweenness Distribution](image)
Limitations

The following caveats should be noted:

1. The cluster analysis was done on my social network. Another person’s social network may have significant differences in how clusters form.
2. The facebook friends network does not reflect a real-life friends network. Maybe connections are in the form of acquaintances or random encounters.
3. This study fails to take into account all my friends that are not on facebook.

Future Work

Some of these topics would be interesting to look at for future development.

1. Develop a directional network (facebook friends only asks for common consensus) where as real-life friends networks have a directional component. Data could be gathered to see who “requested” to be friends on the facebook. It’d be meaningful to see which factors cause a person to have a very high indegree (a lot of people want to become friends with him/her).
2. Developing a weighed network to find where the strong + weak ties are. Data could be gathered on how often messages are sent between nodes to denote the strength of the tie. It’d be meaningful to see which factors will influence a person to develop the strongest relationships.
3. Go in depth to see other factors such as parties, barcrawls, internships, form the most connections. Data could be gathered on the timing of the friends request (a facebook planned party leads to a friend request. It’d be meaningful to see which factors have the highest influence on creating friendship networks.
4. Develop a global network including friends from other schools. Data could be gathered using the same method in this study. It’d be meaningful to see what factors form global friendship networks.

Conclusion

This study showed the school of study was the biggest factor in forming clustering friendship networks on facebook. A possible explanation is that people spend the most time meeting other people in a learning environment such as going to classes, doing homework, joining clubs. Year in school and international status were not strong influencers of forming clusters. With the limitations and future possibilities in mind, there is much to be explored in the field of social networks.