# Social Network Analysis Using Gephi

## Introduction

This project explored the dataset "CLASS OF 1880/81: friendship network of a German boys' school class from 1880/1881" dataset (2014). "It's based on the probably first ever primarily collected social network dataset, assembled by the primary school teacher Johannes Delitsch".

A Force Atlas was created to determine the layout of the graph, then modified to explore different categories of students (based on three variables: repeater, sweetsgiver, handicapped).

The larger nodes in graph 1 have the most degrees of connection. Some student had no or very low degrees of connection. Graph 2 was created to explore different categories of students and discovered that the handicapped individuals in this class have the least degrees and need to be better integrated.

*Graph 1 – Force Atlas, Size Denotes Number of Degrees*

Color designates groups. Larger nodes mean more degrees of connection.
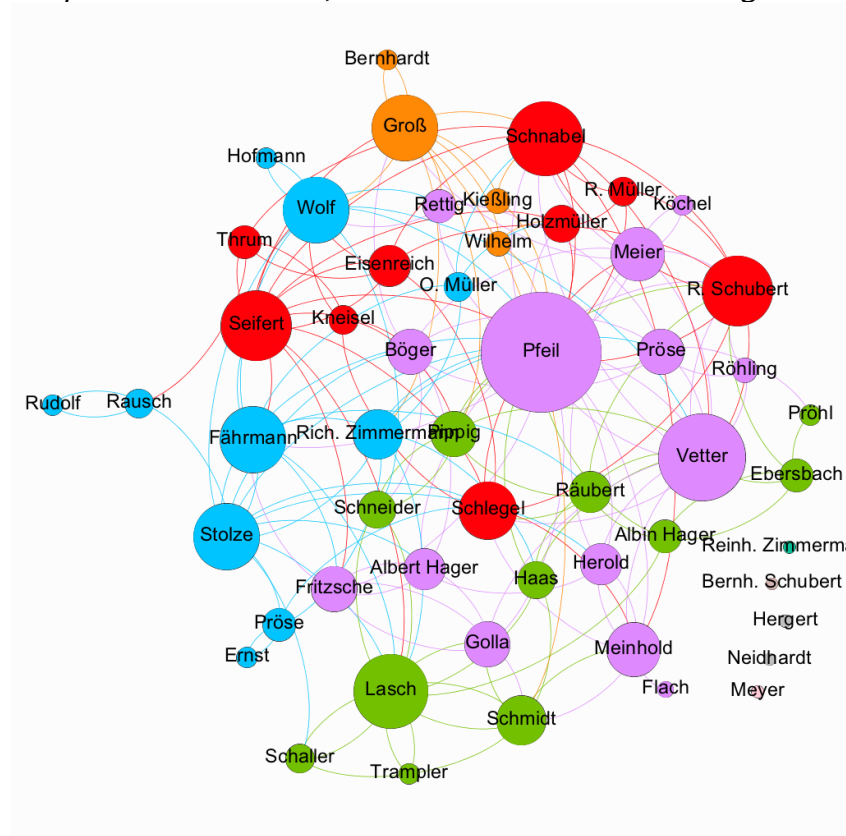
Most members of a group speak to some but not all of the other members within that group as well as a few members from other groups.

*Graph 2 - Force Atlas with Handicapped Ranked by Size*
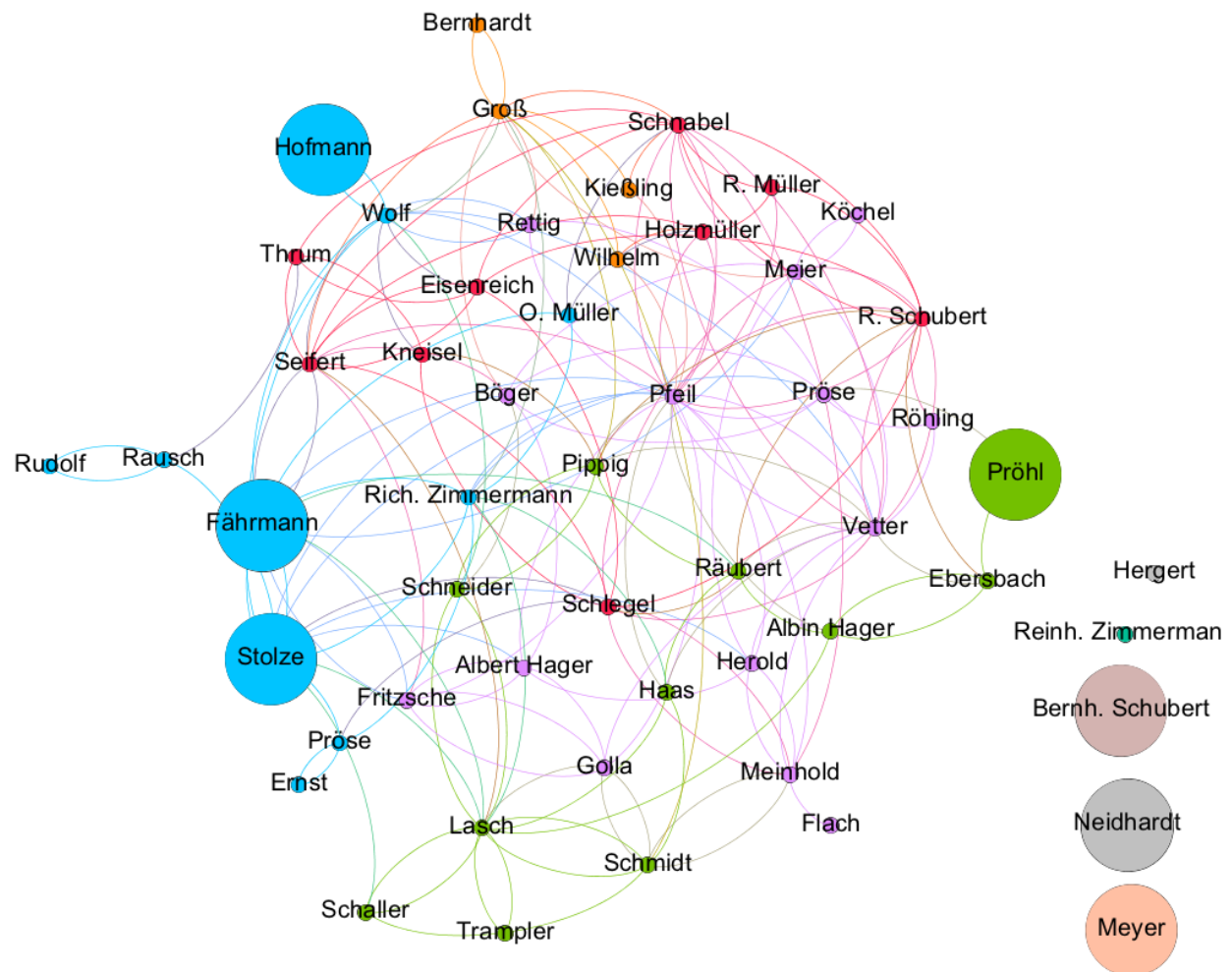


Size was used to locate the handicapped students while keeping color to designate their group.

A Yifan Hu was ranked by Betweeness Centrality to identify which students could most rapidly diffuse information in the network. The nodes with the highest betweeness centrality also tended to have the highest number of degrees.

*Graph 3 - Yifan Hu – Betweeness Centrality*
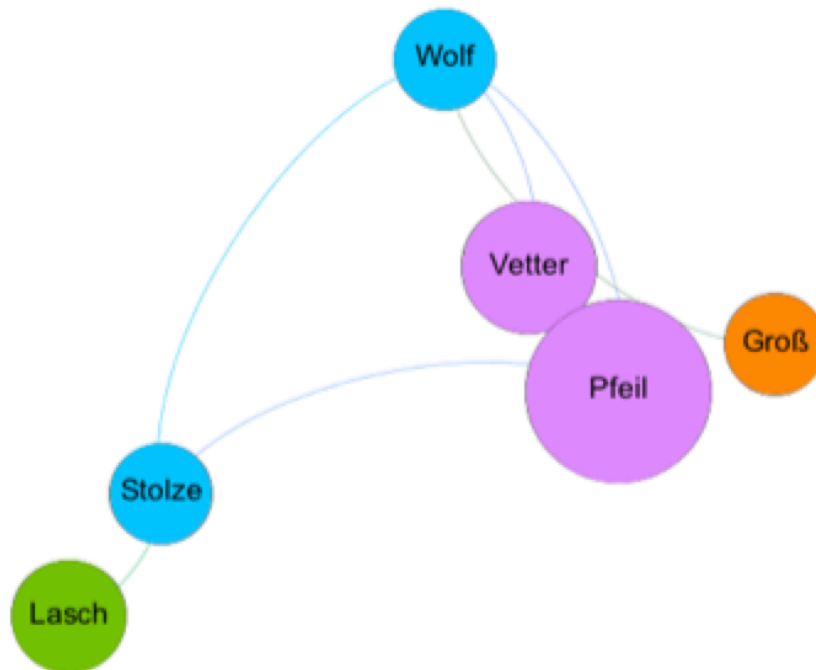


## Conclusions

There are 3 variables provided by the dataset: repeater, handicapped, and sweetsgiver. There were 10 groups identified, 5 of the groups were isolates (3 of which were handicapped). Only the handicapped variable helped to explain any of the groupings.

Handicapped students had fewer contacts than others. 3 of the 7 handicapped students were isolates. And the other handicapped students had a smaller than average degree values.

There must be some other feature, perhaps common hobbies, friendship, or the recess activity they engage in that unites these groups.

The Force Atlas ranked by degree identifies the students with the highest degrees. Wolf, Gross, and Stolze have 13 degrees, Lasch 15, Vetter 18, and Pfeil 26. The Yifan Hu ranked by betweeness centrality determines that the students with the shortest paths to other nodes are the same as those having the highest degrees. These students are the center of the social network and can easily disseminate information to most others within the network.

Most members of groups speak to some but not all of the other members within their group as well as a few members of other groups. It is a small network that has good ability to share information among nodes.

From the information on hand, what unifies each group cannot be determined. However, it is apparent who is at the center of the network and that handicapped students need to be better integrated into the network.

## Appendix

## Force Atlas

A Force Atlas was run with Repulsion Strength set to 700 and Adjust by Sizes checked on to avoid overlap nodes. It is composed of 53 nodes and 179 edges.

It has 10 different communities. Communities 2, 3, 7, and 9 comprise 26.42%, 20.75%, 18.87%, and 16.98% of the nodes. Roughly 80% of the nodes are described by these 4 communities.

The communities were partitioned by different colors. The degree was ranked by size to determine which communities had the most and least connections in the network.

Graph:
- 10 communities (2,3,7, and 9 describe 80% of the nodes)
- Diameter: 5
- Average Path Length: 2.352
- Average Degree Value: 3.377
- Number of Distinct Components: 6
- Modularity: 0.259
- Diameter: 5

The diameter of 5 informs us it is a network of medium complexity. Information flows easily throughout most of the network, though the nodes have a small number of linkages on average (3.377). The network is highly interconnected with good community structure (modularity of 0.259).

From the force atlas we can visually identify the students with the highest degrees. Wolf, Gross, and Stolze have 13 degrees, Lasch 15, Vetter 18, and Pfeil 26.

It also allowed for identification of the 5 isolates, communities 0, 1, 4, 6, 8 who are Reinh. Zimmermann, Hergert, Meyer, Bernh. Schubert, and Neidhardt.

There are only 3 columns to identify these individuals: name, whether they are repeating the grade (repeater), and handicapped. There is one variable "sweetsgiver" that is found under the attribute button but not in the data on the Data Laboratory. There is only one student, Lasch, who falls into this category. He does have 15 friends, and this might be why he has so many friends.

There are 7 handicapped individuals spread out among 5 groups. 3 are together in group 9 and Fährman and Stoltz talk to one another and others in that group (9). Hoffman is also in group 9 but only has one friend (who is not handicapped).

Prohl is in group three and has 2 friends.

While four of the handicapped individuals are spread out among the groups, three of the handicapped boys are isolates in groups: 0, 4, 8. This is a high ratio of handicapped who do not have friends. The school could do a better job of integrating these students into the classrooms and with their classmates.

## Yifan Hu Graph

The Yifan Hu graph visually clustered the individuals into their groups better than the Force Atlas.

We can see the individual groups better not just by color but are grouped together better than the Force Atlas.

*Graph 3 - Yifan Hu Graph*
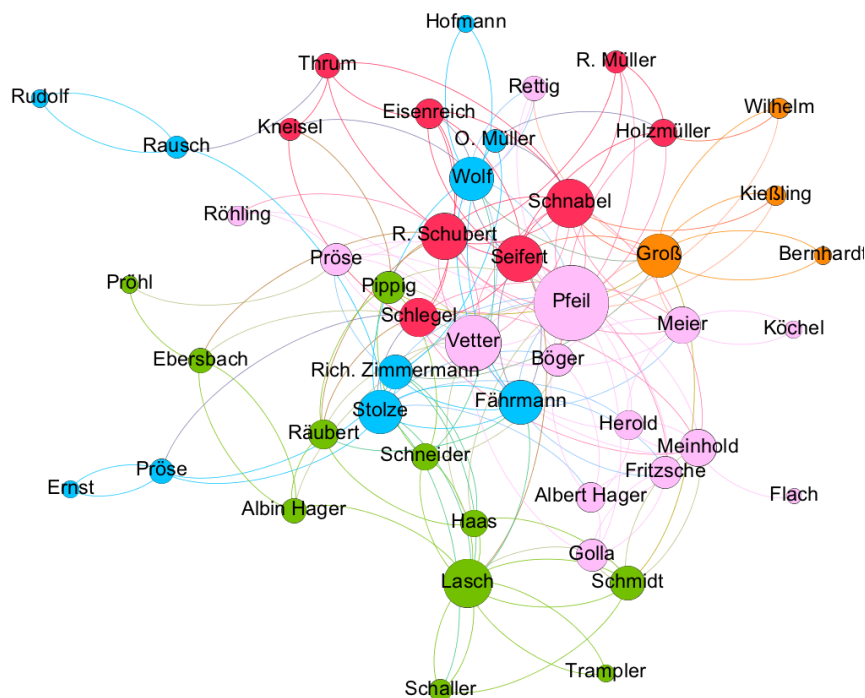
The betweeness centrality was filtered to determine which nodes had the shortest paths to other nodes and could rapidly diffuse information. These are Wolf, Gross, Pfeil, Vetter, Stolze, and Lasch. These are also the ones that are ranked highest in degrees.

## References

GEXF file. CLASS OF 1880/81: friendship network of a German boys' school class from 1880/1881. It's based on the probably first ever primarily collected social network dataset, assembled by the primary school teacher Johannes Delitsch. The data was reanalyzed and compiled for the article: Heidler, R., Gamper, M., Herz, A., Eßer, F. (2014): Relationship patterns in the 19th century: The friendship network in a German boys' school class from 1880 to 1881 revisited. Social Networks 13: 1--13. https://github.com/gephi/gephi/wiki/Datasets