

**SOCIAL NETWORK ANALYSIS ON SCHOOL HEADS AND
SECONDARY SCHOOLS IN THE GOOD SCHOOL GUIDE
22ND EDITION**

Ingrid Lam, supervised by Dr Sol Gamsu



Introduction

Private schools have existed since 597AD as a crucial factor contributing to England's economy and society (Green et al 2010). Private schools differ from state schools in their funding and administration as private schools are financially supported through their students' tuition and independently administered while state schools are supported through public funding and are administered by the Government instead (School Advisor).

Parents justify the personal expenses private schools in how their children can benefit in a private school environment. From more and better consumption services such as sports facilities, music, and cultural goods, to academic services such as higher qualifications and grades, to even non-academic outcomes with examples being private schools offering a better environment to nurture 'soft' skills like better communication and leadership, as well as better networking opportunities with elites ('old boys' or 'old girls' network), the high costs of private schools can be interpreted as an investment for their children (Green et al 2010).

This obviously advantages private school students in future employment and career progression as they are much more likely to obtain high earning, with high social value jobs compared to state school students (Green et al 2010). Because the wealthy can simply pay for access to better educational quality, opportunities, and connections than the middle and low income simply cannot afford to do so, private schools hold high prestige in society and sustains educational inequality, which contributes the growing class divide in England. Educational inequality is defined as the unequal distribution of educational resources such as teaching quality, technology, access to academic opportunities, funding etc. (School Advisor). The comparison between private schools and selective elite state schools in particular should be compared because both types are similar in being selective and holding prestige in society while differing in the requirements to gain entry.

To better study the factors that can contribute to educational inequality, reductionistically investigating factors within the school structure and how it interacts with each other holds merit. Within school leadership, the role most important to the management of schools is with the School Head. A school head is responsible for strategic and budgetary planning, staffing, and most importantly the teaching programme (Trinity College Dublin). They are central to school improvement and are second in their influence on student outcomes. Though indirectly, school heads heavily influence student outcomes in the individual, organisational, policy contexts (Day et al 2014). They are integral in the construction of the school environment. Given the importance of the role, the person that fills this role and their background becomes significant as it reveals not only the technical requirements, but also the possible social requirements and social capital needed to succeed in such a role.

To accomplish this, a type of analysis method called social network analysis can be used. Social network analysis is a distinct approach that has its foundational disciplines based on mathematics, anthropology, sociology, and psychology. It uses a relational approach to understand the world and different systems associated whether this be biological, social, informational etc. (Douglas, 2015).

On a social network, it can be organised into nodes which act as the person who is represented, and edges which are the connections between the nodes. Because humans exist in intertwined networks, social network analysis can be used to find these connections as well as reveal the underlying social processes and patterns for greater understanding of how humans interact with each other and what and how social phenomena occur (Douglas, 2015).

As such, the three core aims of social network analysis is to visualise networks in way that can be instinctively understood even by laymen, to be able to explain specific characteristics of the social network created such as the nodes, edges, and subgroups should there be any, and to construct statistical and mathematical models based on the social network created (Douglas, 2015).

My research aims to create and begin introductory analysis in social network of school heads of private schools and elite state schools by using social network analysis to identify differences and theorize how these differences can lead to educational inequality.

Method

Information was obtained from the Good Schools Guide 22nd Edition. Schools were filtered to only include schools approved by the Good Schools Guide, secondary schools that had attendance until 18, and excluded schools catered for students with special needs or learning difficulties. With each school, information about the region, and the school type (State or Independent) was recorded. The Headteacher's name, history about previous school(s) they worked in, and their corresponding School role (up to the most recent six schools), University attended for undergraduate, College for undergraduate University, Subject studied for undergraduate university, University attended for Postgraduate Certification in Education (PGCE), University attended for postgraduate (Masters or PhD), Subject studied for postgraduate, what Secondary school they attended, the type of Secondary school attended (State or Independent), their Mother's occupation, their Father occupation, and their Partner's occupation were also recorded based on the information available in the Good Schools Guide 22nd Edition. Any information that was not available in the Good Schools Guide 22nd Edition was not further researched. This was completed on Excel.

Once data collection was finished, the dataset was cleaned so that alternate names of the same institution, positions and subject were consolidated into one. Initials of institutions were used wherever possible for ease of reading and better comprehension (e.g., Kings College London would be converted into KCL).

Data analysis was conducted on R. Using the tidyverse package, data was first converted from wide form to long form into a bipartite edge list. Wide form refers to a table in which the data flows horizontally where the same variable is measured across different columns. Long form refers to a table in which the data flows vertically where each variable has its own column. The reason to convert data from wide form (the format data was encoded) to long form is so emphasis can be put on the Headteacher and the connection to the school. This allows easier reading of the data and better shows the distinct types of

variables in the data. In this instance, the variables would be the Headteacher and the number and order of schools.

Only data about the school's name, the Headteacher's name, history of the previous school(s) the Headteacher worked in, and their corresponding School role (up to the most recent six schools) were selected for initial analysis. *Fig 1.* shows an example of code that can be used to build a bipartite edge list and the results.

```
> library(tidyverse)
> gapminder_clean<-read.csv("REAL_GSG_Heads_Database_for_Ingrid.csv",na="")
> heads_data%>%pivot_longer(col=starts_with("PrevSch"),names_to = "School",names_prefix = "PrevSch",values_to = "School Name",values_drop_na = TRUE)
# A tibble: 4,599 x 3
  Headteacher School `School Name`
  <chr>        <chr>   <chr>
1 Michael Windsor 0 "Abingdon School"
2 Michael Windsor 1 "Reading Blue Coat School"
3 Michael Windsor 2 "RGS Guildford"
4 Michael Windsor 3 "King's College School"
5 Michael Windsor 4 ""
6 Michael Windsor 5 ""
7 Michael Windsor 6 ""
8 Nicholas John 0 "Acland Burghley"
9 Nicholas John 1 "St Paul's Way Trust School"
10 Nicholas John 2 "Greate Yarmouth High School"
# ... with 4,589 more rows
# i Use `print(n = ...)` to see more rows
> |
```

Fig.1

'Headteacher' is the column for the names of Headteachers present in the Good Schools Guide 22nd Edition. 'School' is the column for the number of schools with '0' being the current school, '1' being the most recent school the Headteacher has worked in, and '6' being the earliest school the Headteacher has worked in. 'School Name' is the column for the names of the school.

To then conduct one the core elements of social network analysis, which is to visualise the network, the igraph package on R was used to plot the connections between schools and headteachers. *Fig.2* shows an example of code that could be used to create a social network.

```
1 heads_data<-read.csv("REAL_GSG_Heads_Database_for_Ingrid.csv")
2 str(heads_data)
3 library(tidyverse)
4 gapminder_clean<-read.csv("REAL_GSG_Heads_Database_for_Ingrid.csv",na="")
5 heads_data%>%pivot_longer(col=starts_with("PrevSch"),names_to = "School",names_prefix = "PrevSch",values_to = "School Name",values_drop_na = TRUE)
6 edge_list<-heads_data%>%pivot_longer(col=starts_with("PrevSch"),names_to = "School",names_prefix = "PrevSch",values_to = "School Name",values_drop_na = TRUE)
7 library(igraph)
8 g<-graph.data.frame(edge_list,directed = FALSE)
9 plot(g,vertex.label.colour=c("black"),vertex.label.cex=c(0.5),vertex.size=2,layout=layout_with_lgl(g))
10 install.packages(igraph)
```

Fig.2

Out of all the possible layouts, the *layout_with_lgl*, which stands for 'layout with large graph layout' was chosen due to the large volume of data collected (list of 657). This code was created by Alex Adai (McWhite, 2016). *Fig.4* shows the results of the code and visualises the nodes and edges of headteachers and the schools they have previously worked with.

Once visualised, introductory analysis was conducted by running community detection algorithms. The 'community' in community detection algorithms refers to nodes that cluster together with smaller edges than compared to other nodes in the same graph (Joshi, 2020). Therefore, a community detection algorithm scans and analyses the social network. It does this by splitting the social network into multiple different communities in two methods: through a divisive or agglomerative approach.

A divisive approach begins with a full graph with both nodes and edges and removes edges also starting from the 'strongest'. An agglomerative approach differs in that it begins with an empty graph with only the nodes present first before adding edges on a 'strength' basis, starting from the 'strongest' to the 'weakest'. (Joshi, 2020).

The Newman-Girven Edge-Betweenness algorithm was used with results seen in *Fig.5*. This community detection algorithm works by first calculating the edge betweenness centrality for every edge on the graph. This means measuring the number of vertices a node may have. Having many vertices (many edges connected to one node) may signify that that node has a lot of influence on that network and can be in the center of a community or even in the center of the graph. The Newman-Girven Edge-Betweenness then removes the edge with the highest betweenness centrality, which is an example of a divisive approach. The algorithm then calculates the betweenness centrality for the remaining edges and the process repeats until there are no more edges left (*Girvan-Newman Algorithm | NetworkX Guide*, n.d.). A dendrogram can also be produced based on the data to analyse hierarchical clustering, as seen in *Fig.6*.

The Label Propagation algorithm, which is an agglomerative approach, was also used as seen on *Fig.7*. How this algorithm works is that each node has a unique label, and these labels propagate throughout the network by having randomly selected nodes adopt the label of the majority label of its neighbors. This stops when every node has the majority label of its neighbors (Smith et al., 2020).

An example of the code used to program different community detection algorithms can be seen on *Fig.3*. 'edge_betweenness' and 'edge.betweenness.community' codes for the Newman-Girven Edge-Betweenness algorithm, 'dendPlot' codes for the dendrogram, and 'label.propagation.community' codes for the Label Propagation algorithm.

```
11 library(bipartite)
12 betw<-edge_betweenness(g)
13 ebc<-edge.betweenness.community(g,directed = F)
14 plot(ebc,g,vertex.label.colour=c("black"),vertex.label.cex=c(0.5),vertex.size=2)
15 dendPlot(ebc,mode="hclust",rect=5)
16 lp<-label.propagation.community(g)
17 plot(lp,g)
```

Fig.3

Results

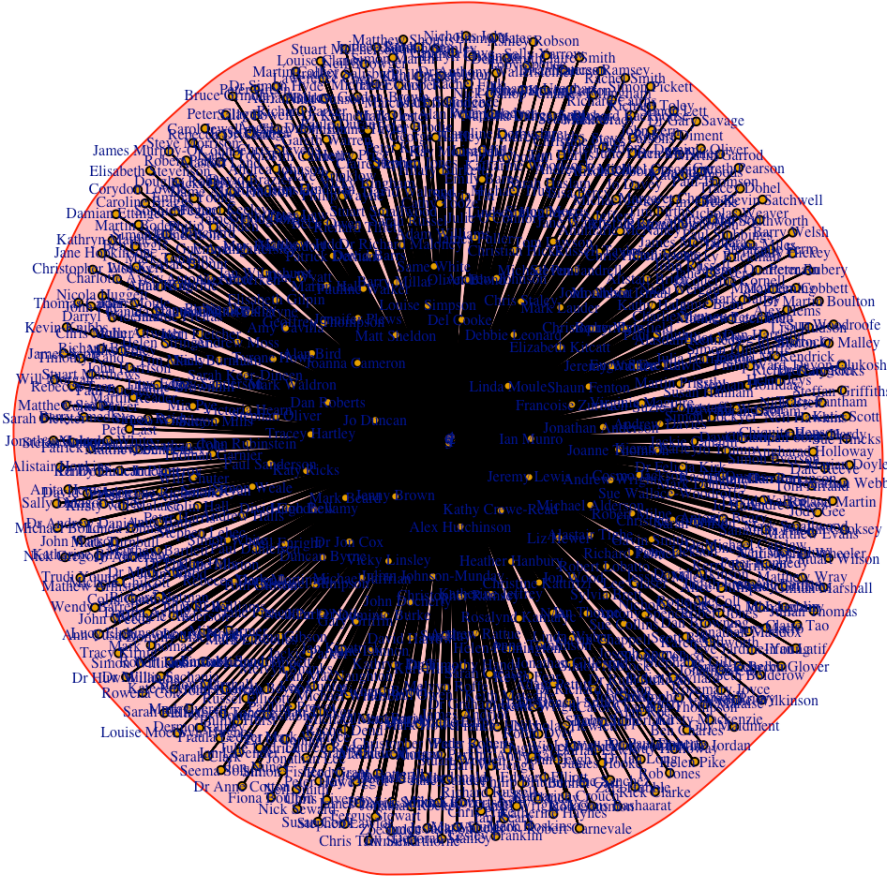


Fig.5

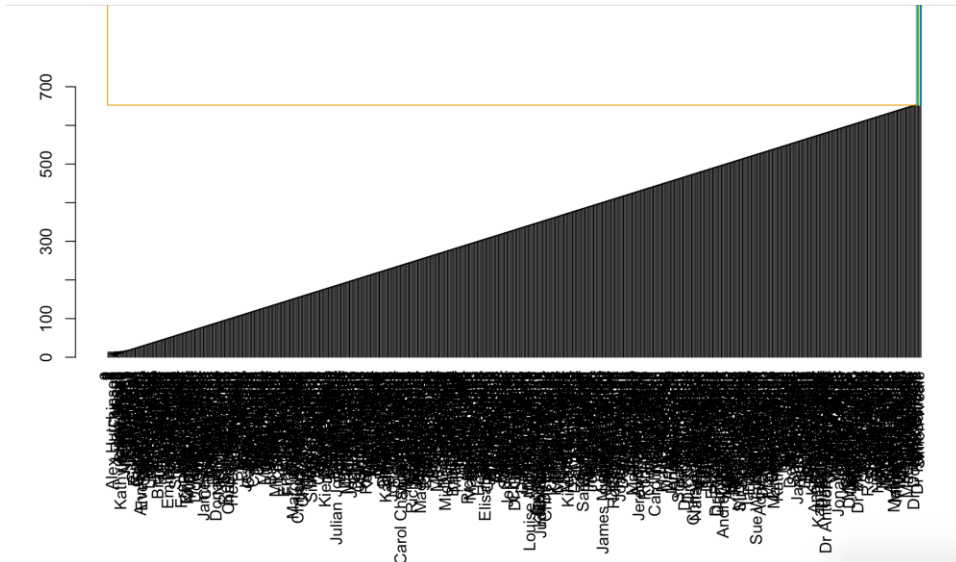


Fig.6

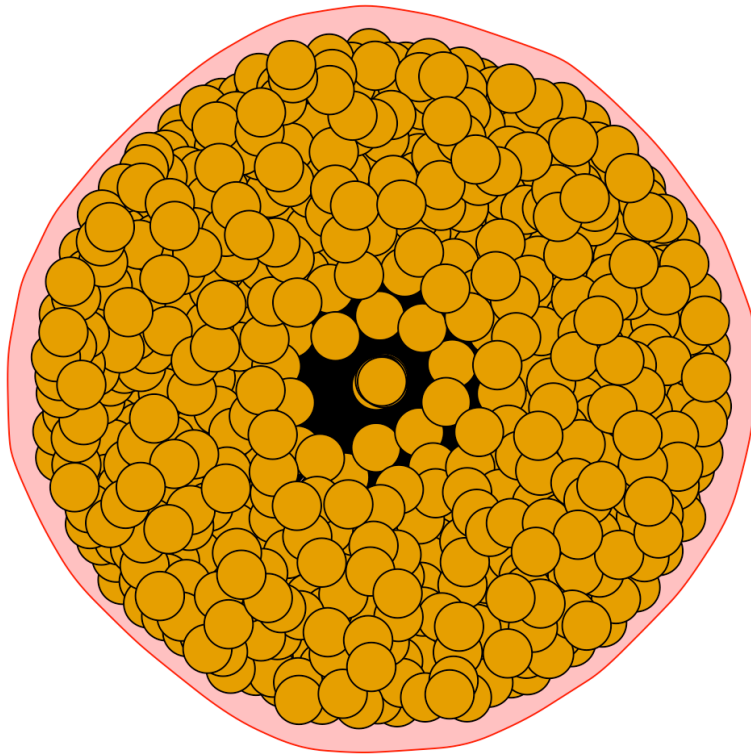


Fig.7

Discussion

Fig.4 shows the social network of headteachers and schools. Its shape is quite circular with many nodes spiking out and a community of headteachers/schools at the very center of the social network. Given the shape of the social network, it can be interpreted that the nodes at the edges of the social network may be headteachers or schools that are not very well established while the community at center is the core of the school/teaching community in UK.

Fig.5 suggests that the social network is one big network, rather than being made up of different communities with edges that connect between them as if this was the case, there would be different coloured circles encircling different communities. As such, this suggests the secondary school education industry is quite intertwined with each other with no distinct communities. *Fig.6* is a dendrogram that has a linear line which differs from the usual shape of a dendrogram. Further research is needed to understand this.

Fig.7 further builds on the discussion that *Fig.4* brings, it highlights the increased edge distance between the majority label of the neighbors in the center of the social network compared to the other majority labels of the neighbors that permeates the rest of the community.

However, it is clear that the discussion above is very introductory and in no means conclusive. Further analysis with R and further research is needed to better understand the data collected and to answer the questions this research brings such as how does the data above link to educational inequality, the specific positions private and state schools are in the social network visualised, how the participant variables of school heads such as whether they themselves were privately or publicly educated contribute to the position they are in the social network and more.

Due to time constraints and the limited understanding given their educational experience thus far of the author, further analysis and understanding cannot be uptaken. This report provides an introductory visualisation into the social network of school heads and private and public schools and encourages further research and communication on the topic of educational inequality in the UK.

References

- Day, C., & Sammons, P. (2014). Successful School Leadership. *Education Development Trust*.
<https://www.educationdevelopmenttrust.com/our-research-and-insights/research/successful-school-leadership-original-2014-publica>
- Douglas, L. (2015). *A User's Guide to Network Analysis in R*. Springer International Publishing.
<https://doi.org/10.1007/978-3-319-23883-8>
- Girvan-Newman algorithm | NetworkX Guide*. (n.d.). Retrieved October 31, 2022, from
<https://networkx.guide/algorithms/community-detection/girvan-newman/>
- Green, F., Machin, S., Murphy, R., & Zhu, Y. (2012). The Changing Economic Advantage from Private Schools. *Economica*, 79(316), 658–679. <https://doi.org/10.1111/j.1468-0335.2011.00908.x>
- Head of School - Role and Responsibilities - Secretary's Office - Trinity College Dublin*. (n.d.). Retrieved November 1, 2022, from <https://www.tcd.ie/Secretary/academic-governance/head-of-school.php>
- Joshi, P. (2020, April 12). Getting Started with Community Detection in Graphs and Networks. *Analytics Vidhya*. <https://www.analyticsvidhya.com/blog/2020/04/community-detection-graphs-networks/>
- McWhite, C. (2016, March 23). *LGL - the Large Graph Layout*. Claire McWhite.
<https://clairemcwhite.github.io/lgl-guide/>
- Smith, N. R., Zivich, P. N., Frerichs, L., Moody, J., & Aiello, A. E. (2020). A guide for choosing community detection algorithms in social network studies: The Question-Alignment approach. *American Journal of Preventive Medicine*, 59(4), 597–605. <https://doi.org/10.1016/j.amepre.2020.04.015>
- The Differences Between Public Schools and Private Schools*. (n.d.). Retrieved November 1, 2022, from <https://schooladvisor.my/school-guide/difference-public-schools-private-schools>