A Network Science Approach to the World of Painters

Northwestern
WEINBERG COLLEGE OF ARTS & SCIENCES

Eliseu Antonio Kloster Filho†

†Department of Physics and Astronomy, Northwestern University, Evanston, IL, USA

Northwestern University

Introduction

Art is one of the oldest forms of human expression. As such, the network formed by artists and their impact throughout history is large and intricate. In this project, we aim to uncover some of its structure and understand how it evolves. To accomplish that, we make use of Network Science tools to analyze a network of influences. We treat painters as nodes and construct directed links from teachers to pupils, from influencers to the influenced, and from leaders to their followers.

Data Collection and Processing

The data used in this project comes from the RKDartists online database, which is maintained by the Netherlands Institute for Art History (RKD). Basic information such as artists' names, birth dates, and places of activity is available as Linked Open Data (LOD) for more than 250 thousand people.

However, the relevant network data (teacher-pupil and influence relationships) is only available as text remarks in the RKD website. Therefore, to collect this data, a custom webcrawler was programmed using Python. The crawler searches the website for target relationship keywords, matches the artists's names to their LOD codes using the RKD API and stores the information in a local database that can be queried using SPARQL language.

Custom SPARQL queries were created to convert samples of the LOD data into a format that can be easily transformed into a networkx graph.

Network Properties

Basic Network Properties were calculated using networkx and Gephi. Statistics are available in **Table 1**. Degree distribution is shown in **Figure 1.**

Table 1: Network Summary

Number of Nodes	13569	Minimum Degree	1
Number of Links	21510	Maximum Degree	172
Average Degree	1.585	% Teacher-Pupil Links	78.6%
Average Path Length	9.103	% Influencer-Influenced Links	17.8%
Clustering Coefficient	0.02	% Leader-Follower Links	3.6%
Network Diameter	25		

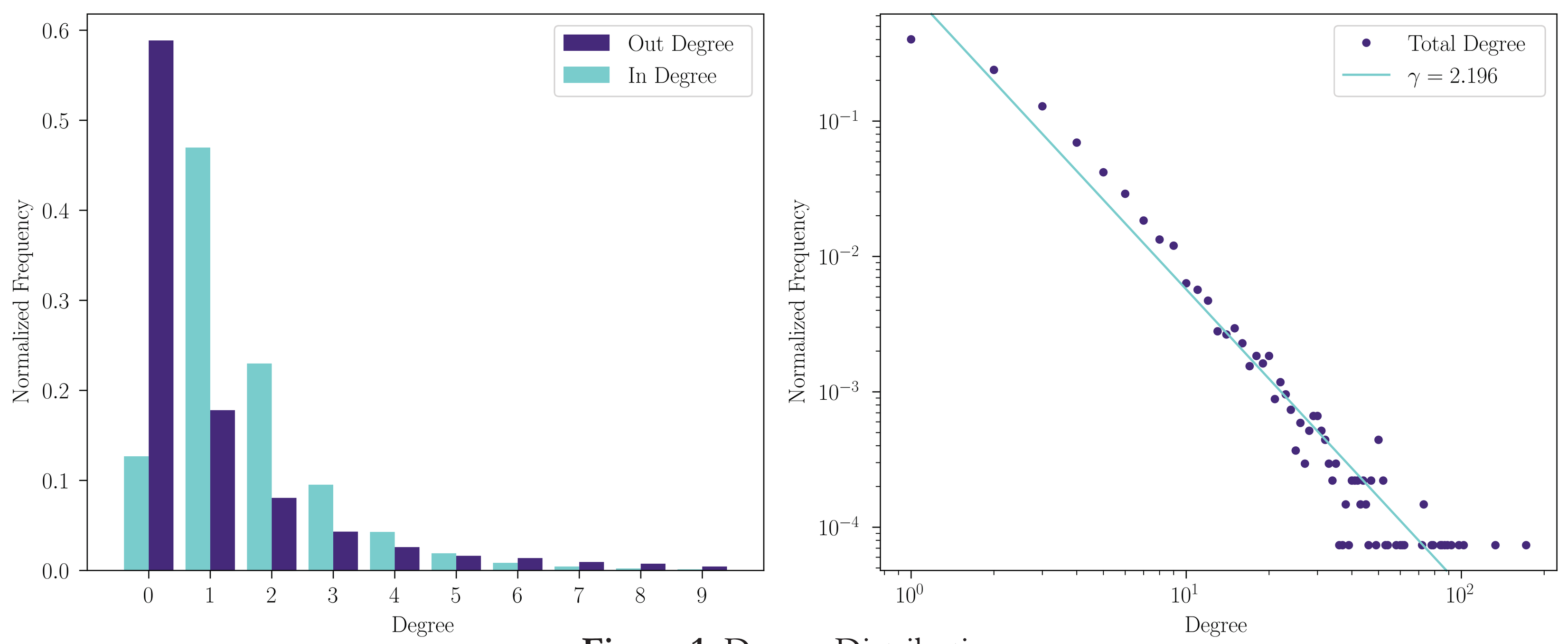

Figure 1: Degree Distribution

Community Detection

The network was imported into Gephi, a network analysis and visualization software. Two community detection algorithms available in Gephi were used: modularity and statistical inference. Modularity peaked at around 0.76 for various choices of resolution, always with more than 30 communities. Statistical inference found 7 communities, with a description length of 18750. The network layout shown in **Figure 2** is colored according to the communities found by the statistical inference algorithm.

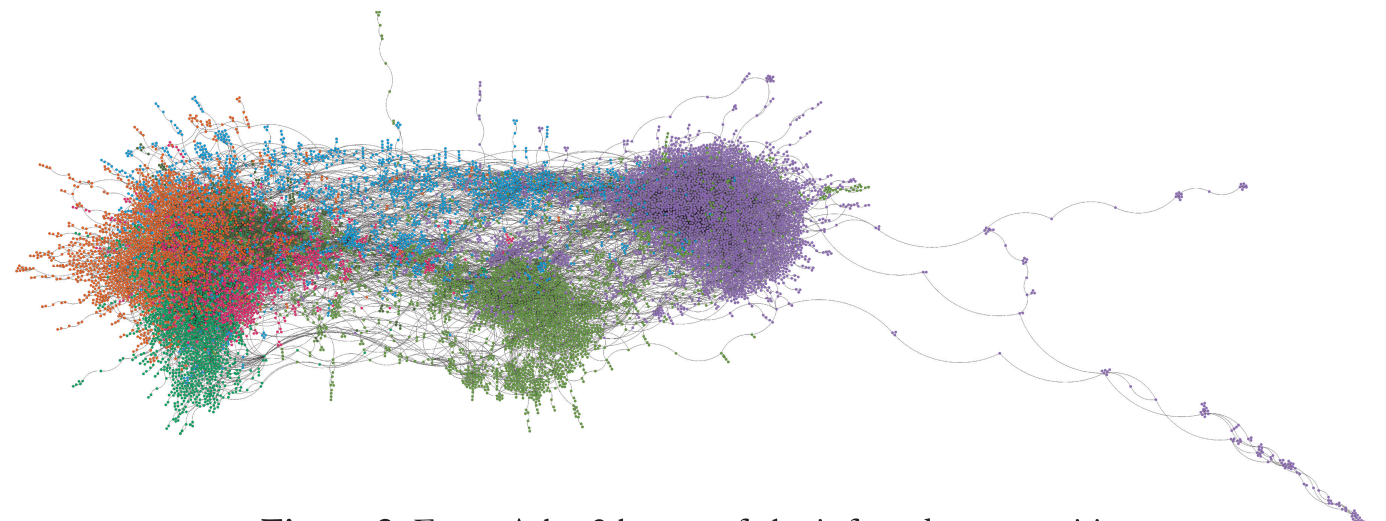

Figure 2: Force Atlas 2 layout of the inferred communities

One reasonable hypothesis is that communities are related to geographical proximity, nationality, or language. Since the database lists place of activity for each artist, we tested this idea. By plotting the nodes at their physical location in **Figure 3** , we see that nodes in the same community tend to be closer geographically, and that geopolitical barriers agree with the regions of community separation.

Visualization

In this poster, the nodes are colored according to the communities detected by statistical inference. **Figure 2** uses the Force Atlas 2 algorithm for the layout. **Figure 3** uses geographical positions for the layout. The graph was zoomed in around Europe because that's where most of the data is. Since we know the approximate period of activity of each artist, it's possible to create a video illustrating the network evolution. After this was done, we found that each of the communities has a time period in which it's more dominant.

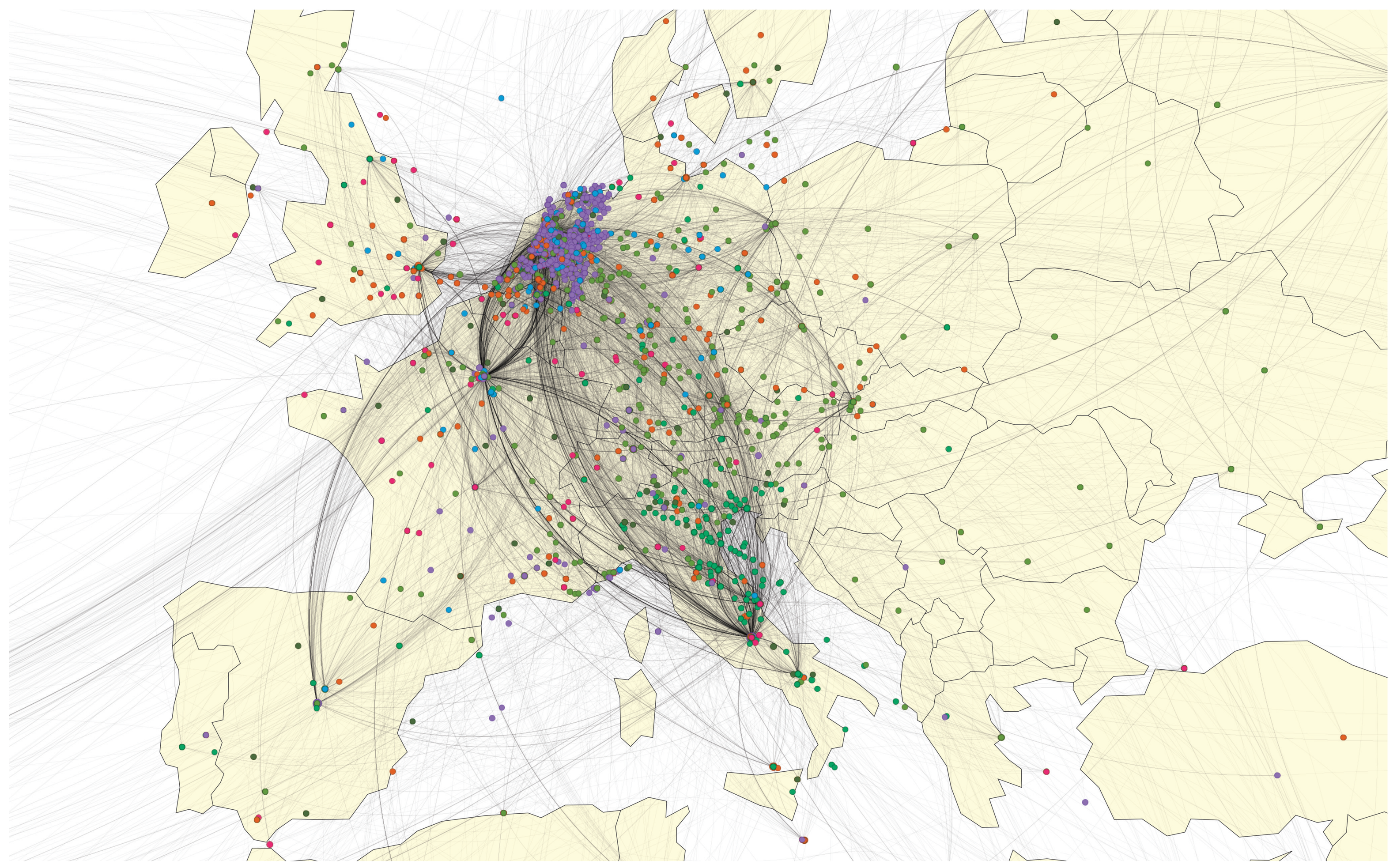

Figure 3: Geographical layout of the inferred communities

Quantifying and Predicting Influence

Scientific citation networks are similar to this network in the sense that they are networks of influence. To test if they can be similarly described , we fit the model used by Wang et al [2] to quantify long-term scientific impact, shown in **Equation 1**. **c(t)** is the cumulative out degree , **m** is a global parameter, **Φ** is the cumulative normal distribution, and the other variables are fit to each node. We find that the higher degree nodes can be described by this model. **Figure 4** shows the time evolution of out degree (purple) compared to the model's prediction (blue) for a high-degree node (Jacques-Louis David).

$$c(t) = m\left(e^{\lambda\phi\left(\frac{\ln t-\mu}{\sigma}\right)}-1\right) \quad (1)$$

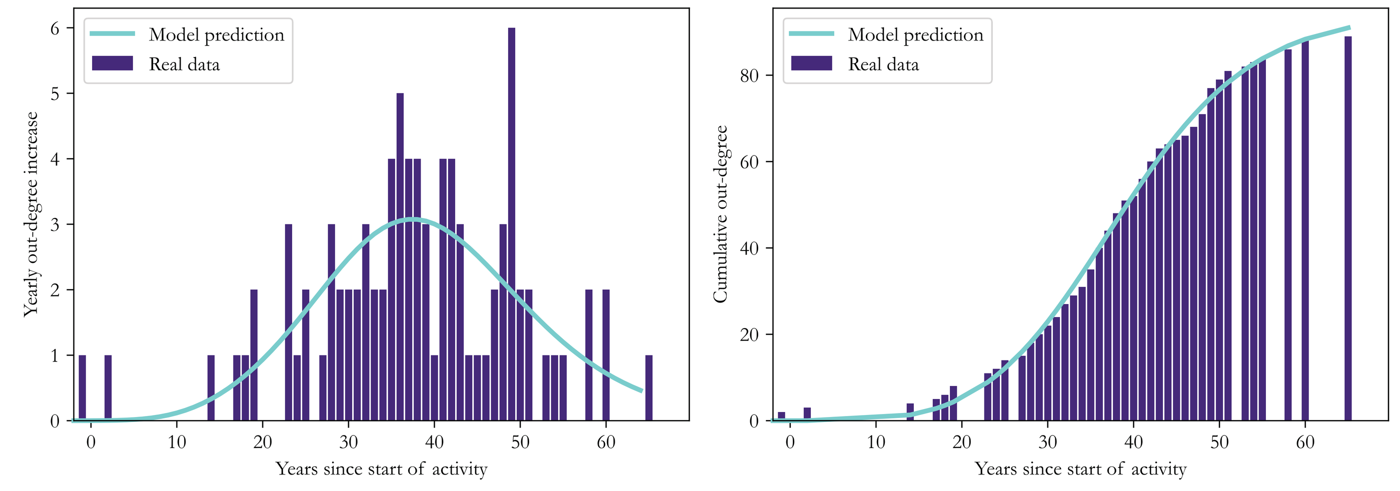

Figure 4: Degree evolution prediction for a high-degree node (Jacques-Lous David)

Discussion

We find evidence for scale-freenes given the observed degree distribution and the nature of the network. We also have reasons to believe that communities of artists are related to their geographical locations, as seen in **Figure 3**. On the other hand, there are biases affecting these results: most of the data comes from times before the emergence of fast communication technologies and globalization, which means they might not be good predictors of a contemporary network. Because of the source of the data, we also observe a much higher amount of nodes in the vicinities of the Netherlands

We found that the high-degree nodes (> 10) can be described by the same model used in citation networks. We can therefore rank them by their ultimate impact **c(∞)=m(exp(λ)-1)**. However, we must be careful to interpret this correctly: a significant portion of the artists predicted to be highly influential are actually just teachers in various art schools, so their influence is not necessarily due to pure talent, as one might desire to find.

Although available, we did not analyze data on techniques used, common subjects, institutions, and awards. Further work could be done in that regard.

References

[1] Wang, Dashun, Chaoming Song, and Albert-László Barabási. 2013. "Quantifying Long-Term Scientific Impact." Science 342 (6154): 127–132.

[2] Barabási, Albert-László. 2016. Network Science. 1st ed. Cambridge University Press.

[3] RKD – Netherlands Institute for Art History