culturegraphy

Kim Albrecht
Potsdam University of Applied Sciences
Kiepenheuerallee 5
Potsdam, Germany
me@kimalbrecht.com

Marian Dörk
Potsdam University of Applied Sciences
Kiepenheuerallee 5
Potsdam, Germany
doerk@fh-potsdam.de

Boris Müller
Potsdam University of Applied Sciences
Kiepenheuerallee 5
Potsdam, Germany
boris.mueller@fh-potsdam.de

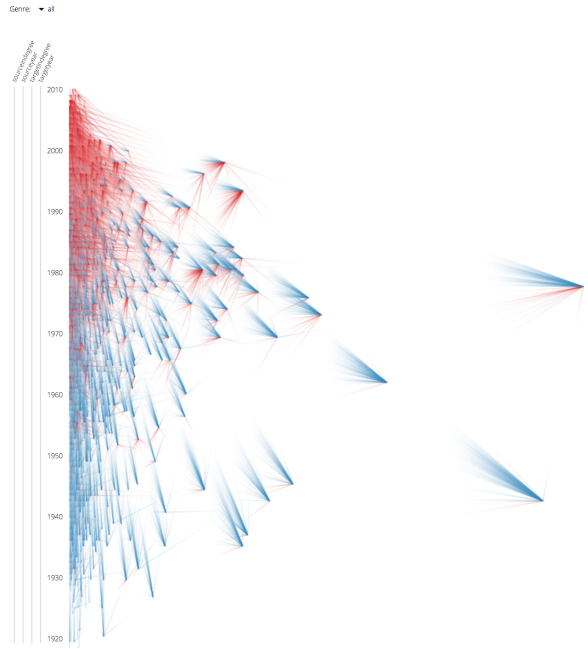

Figure 1: Movies by connectedness

1. ABSTRACT

Culturegraphy visualizes the exchange of cultural information over time also known as memes. Treating cultural works as nodes and influences as directed edges, the visualization of these cultural networks can provide new insights into the rich interconnections of cultural development. The graphics represent complex relationships of movie references from IMDB (Internet Movie Database) and influences between popular individuals from Wikipedia. All findings were made in a process which involved network scientists, a media theorist, and a sociologist. The role that visualization can play in bridging scientific communities was central to this work. In this sense, the resulting visualizations were process involving to bring researchers from different disciplines together. Today physicists through the study of networks ask similar questions as media theorists or sociologists with very different techniques and methods. Visualization can serve as a common language that brings fields together, shows differences, but also has its own idiosyncratic views.

2. INTRODUCTION

Throughout this project, we investigate how to visually explore the dynamics of memes [5, 6, 2] traveling through cultural history and present the first findings that we were able to make by visualizing this model of culture. We ask ourselves what happens when we look at culture as a *superorganism* [10] in which knowledge is copied, combined and transformed from person to person as well as their artefacts over time in a network of interactions. What can we learn from visual representations of this model of culture? The interactive visualizations are explorable at culturegraphy.com. The data is based on two datasets. One from Wikipedia on influences between well known artists, writers, philosophers, musicians and others and from the IMDB on movie citations. We were able to make findings on the macro level of the graphics and on the level of each individual movie through the help of experts in different fields and a variety of different visualization techniques that represent the network on different levels of abstraction. This position paper gives an overview of the process and describes particular findings that we gained.

3. RELATED WORK

Culturegraphy relates to several visualization techniques intended to represent network structure in new visual forms. The most common network visualization techniques do not utilize the plane efficiently. In linear representations only one axis is used to plot data. Same can be said about circular graphs where the nodes are ordered radially rather than linearly. In force-directed graphs where the nodes are positioned by physical forces the plane is not used in this case to represent data spatially at all. In 2006, M. Wattenberg developed a visualization technique called PivotGraph [16]. The technique uses a very simple but also quite restricted approach of placing the nodes on a grid structure by aggregating nodes by attributes. A. Aris and B. Shneiderman created a visualization technique for networks called semantic substrates [1] that used some of the ideas from PivotGraph. It

is a spatial template for a network, where nodes are grouped into regions and laid out within each region according to one or more node attributes. The main contribution is the idea of different 'substrates' in which the nodes are placed. Since then multiple projects picked up the idea to display networks in other spatial formations. M. Dörk, S. Carpendale, and C. Williamson used a multi-dimensional scaling algorithm to calculate implicit similarities between items and map them on the plane in their project 'EdgeMaps' [8]. Jan Willem Tulp created an interactive version of 'The Flavor Connection Network' [15]. Tulp ordered the nodes in the visualization by the number of connections on the y Axis and grouped the network into categories on the x Axis. The approach is similar to semantic substrates but by restricting the regions to the x Axis the graphic stays much clearer and easier to read.

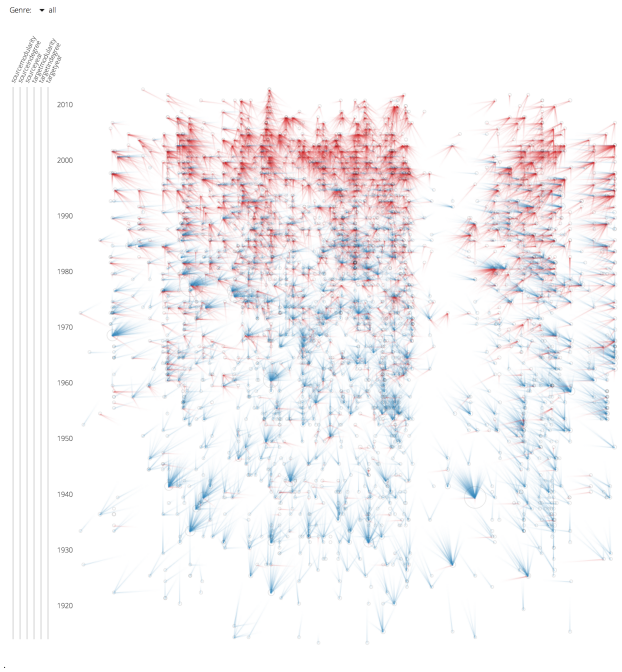

Figure 2: Movies by community

4. DESIGN

Building on network visualizations that arrange nodes by attribute, we introduce 'culture.graphs', multiple visualization techniques that integrate the representation of temporal dynamics with graph-theoretical metrics.

4.1 By connectedness

The first version of 'culture.graphs' was an arrangement of the nodes by year and by indegree. Indegree is the number of movies who referenced the selected movie or the connectedness of that movie. In force-directed layouts the highly connected movies are placed in the center of the plane. Surrounded by all the movies they are connected to. In our arrangement the highly connected movies stand out and become the outliers on the right of the graphic that stand alone while the low connected movies create a dense area on the left of the plane. This creates a shift in perspective on what is important in the representation. While force-directed graphs in the best case represent different clusters

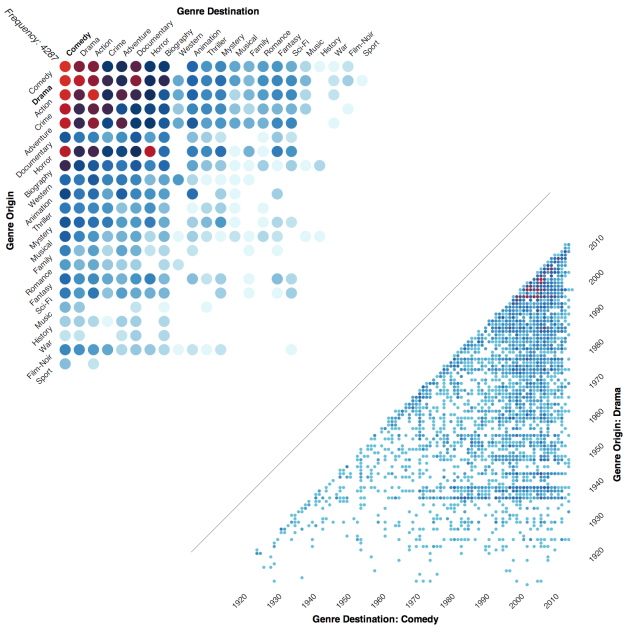

Figure 3: Movies by genres. The visualization over time shows memes coming from 'Drama' adapted by 'Comedy'.

in the network, the ordering by degree shows the connection between time and the popularity between individual movies.

4.2 By communities

The decision to place nodes by their connectedness leads to interesting results and makes it easy to see the highly connected nodes in the network. However, high connectivity is not the only interesting aspect of a network. One thing that force-directed networks are particularly useful for is arranging nodes into groups within the network. They do so based on physical forces on the nodes and edges and by useing both axis of the plane. We tried to find an algorithm that orders nodes into groups on statistical models on one axis so we are able to use the other axis to represent the temporal aspect of the network. We looked at modularity measurements [4], which take links away from the network until it breaks apart the clusters that come up in this process become the individual groups that this algorithm produces. We used this approach and visualized the different modules on the x axis of the graph. While arranging the nodes by modularity is not the most precise method to represent clusters in a network, as a one-dimensional metric it can be used for one axis and always returns the same results.

4.3 By genres

A different visualization technique was used to represent the interconnectedness of different genres in the movie data. While the two previous visualization techniques, by connectedness and by communities are using lines to represent networks, we arranged points in an adjacency matrix to represent idea exchange between different genres. The visualization is based on two matrices in the upper graphic all genres and their exchange is visualized. By clicking on a genre in-

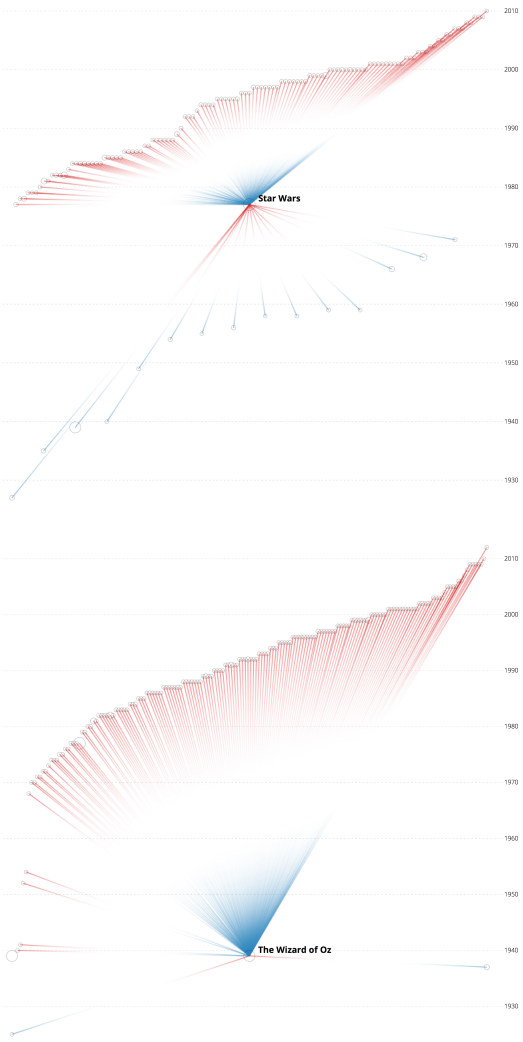

Figure 4: Visualization 'by movies'. The upper graphic shows the citation pattern of 'Star Wars'. In comparison, the citation pattern from 'The Wizzard of Oz'.

teraction a second matrix represents the interaction between the two selected genres over time. This network within a network shows trends over time and makes patterns in the citation behaviour visible.

4.4 By movies

The fourth visualization focuses on the representation of one node at a time and its connections. While the previous three visualizations are making macro connections visible this view lets the viewer explore the connections in a rather 'monadic' approach [11]. The citation pattern over time of individual nodes becomes visible, explorable and comparable in this view.

5. MOVIE REFERENCES

The movie visualizations are based on data from IMDB which is an online movie community with 42 million members. People contribute various information about movies on IMDB, for example, one part is a collection of connections between movies. For the visualization we focused on references among movies and their release year. The dataset contains 119,135 reference connections from 42,571 movies. To be able to responsively visualize the data in a web browser we selected the 3,000 most connected movies by in-degree and their 10,000 connections to each other.

There is a great variety of references among movies that is captured on IMDB. To give a sense of the types of references, here are two examples for the movie 'Star Wars' from 1977: The robot C3PO was modelled after the robot from the movie Metropolis (1927). C3PO used a line from the movie Seven Samurai from 1954: "It seems we are made to suffer. It's our lot in life". But not only Star Wars referenced other movies they also got referenced in many occasions. The line "May the Force be with you." was referenced by movies like Barney Miller: Quo Vadis? (1978), The Big Fix (1978), Sledge Hammer!: They Shoot Hammers, Don't They? (1986), Beverly Hills Cop II (1987) and many others.

The visualization maps the dynamic structure that is created by the cultural practice of copying, transforming, and combining ideas over time. While positioning the nodes of the graph on the y Axis on a time-line from 1900 to 2015 the x Axis has different versions like degree (number of connections), modularity (closeness of the nodes to each other) or genres the movies are ordered into. The different visualizations can be filtered by year, degree, and by the characteristics of each graphic like genre or modularity. These graphics give a summary over the inspiration in the history of cinema from the perspective of the IMDB members.

5.1 Findings

Until now we can present two 'first findings' throughout our research rather than a full developed study. The findings were made though interviews with scientists from different disciplines. We talked with physicists who are specialized in cultural networks, a media theorist and a sociologist. The two visible most dominant patterns that we found though this process are: The rise of the postmodern cinema on the macro level and on the micro level different citation patterns of individual movies.

5.2 Rise of the postmodern cinema

One well visible artefact of the created 'culture.graphs' is the color gradient within the graphics from the bottom in blue to the red top. The red top which starts around the 1980's is known in film studies as the postmodern cinema [3, 9, 7]. An era in which movies strongly cite the style, stories, and scenes of past movies.

5.3 Citation patterns

While the rise of the postmodern cinema is a dynamic process that emerged on the micro level but only exists as a phenomenon on the macro level, there are also patterns that can be found and analyzed on the micro level of each movie. One difference that can be seen between the individual movies is the citation patterns they have. The movie 'The Wizard of Oz' got released in 1939 and for 30 years, until the beginning of the 1970's, the movie nearly did not receive any references. In comparison to Alfred Hitchcock's 'Psycho'

which got released in 1960 and got referenced immediately after its release. But even with the delay of 30 years 'The Wizard of Oz' got more references than any other movie, except Star Wars.

6. CONCLUSIONS

Network visualization can function as a language [12, 13]. A language to describe complexity that holds the capability to connect scientific fields that are usually unrelated to each other [14]. Through this language new findings can be made that combine micro with macro patterns of complex relationships in ways math, statistics or written language could not. The here presented visualizations should inspire interdisciplinary projects that use this new language as a common ground for their communication. But while science searches for answers the graphics are also there to ask questions. What kind of pictures do we draw from culture? Are these reflections of our mental models rather than reflections of reality? And if so, are they not also cultural objects?

7. REFERENCES

[1] A. Aris and B. Shneiderman. Designing semantic substrates for visual network exploration. *Information Visualization*, 6(4):281–300, Nov. 2007.

[2] S. Blackmore. *The Meme Machine*. Oxford University Press, Mar. 2000.

[3] J. K. Bleicher. Zurück in die Zukunft. In *Oberflächenrausch*. Lit Verlag, Jan. 2008.

[4] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 10(1):008, Oct. 2008.

[5] R. Dawkins. *The selfish gene*. May 2006.

[6] D. Dennett. *Darwin's Dangerous Idea*. Sept. 2006.

[7] J. Distelmeyer. Die Tiefe der Oberfläche. In *Oberflächenrausch*. Jan. 2008.

[8] M. Dörk and S. Carpendale. EdgeMaps: Visualizing Explicit and Implicit Relations. pages 1–12, Sept. 2013.

[9] J. Eder. *Oberflächenrausch*. Postmoderne und Postklassik im Kino der 90er Jahre. Lit Verlag, 2008.

[10] C. A. Hidalgo. The value in the links: Networks and the evolution of organizations. *Sage Handbook on Management and Complexity ...*, 2011.

[11] B. Latour, P. Jensen, and T. Venturini. 'The whole is always smaller than its parts'–a digital test of Gabriel Tardes' monads. *The British journal of ...*, 2012.

[12] S. Ortiz. Santiago Ortiz - Visualized. *visualized.com*, Feb. 2014.

[13] F. Samsel. Art-Science-Visualization Collaborations; Examining the Spectrum. In *Proceedings of the IEEE VIS Arts Program (VISAP)*, 2013.

[14] S. L. Star and J. R. Griesemer. Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social studies of science*, 19(3):387–420, Aug. 1989.

[15] J. W. Tulp. The Flavor Connection [Interactive]. *scientificamerican.com*.

[16] M. Wattenberg. Visual Exploration of Multivariate Graphs. pages 1–9, Jan. 2006.