Anime and Production house analysis using AniList

Bence Gattyan[s204773] and Gokul Desu[s242580]

Animaiton is a relatively new art-form as far as art goes, yet it has been rapidly evolving in the past century. Specifically, Japanese animation (Anime) has been constructing expansive fictional universes which have a multitude of characters and are shaped by complex production networks, yet this network is rarely ever examined. In this paper, we use AniList data to build two different graphs. One is a production network which captures how personnel involved in the production process collaborate across titles, and the other is a character network which shows how characters interact and relate to each other.

We analyze community structure, centrality among other things and try to derive meaningful insights based on these networks. Our results reveal modular but overlapping clusters which highlight non obvious structural hubs on both sides and show industrial connections do not simply mirror on screen communities.Together these findings offer a general framework which is needed to connect the narrative and studio structure in the current anime production field and show how graph based analysis can reveal the underlying structure of the world of anime.

Anime | Network science | Production studios | Fictional world | Community detection | Sentiment analysis | Visualization based analysis

Anime is a form of Japanese animation which is characterized by serialized storytelling, vast network of voice actors and a specialized production pipeline which has a multitude of roles present such as key animators, directors, producers. These series (different anime) generally span many episodes and genres, creating fictional worlds with various characters which are a collaborative effort of studios and the staff. Such a structure makes anime a valuable setting for network science. But most of the work has been focused on Western television, movies and books, leaving the connection between on-screen relationships and behind-the-scenes production largely unexplored.

This project aims to address the aforementioned gap using data we procured from AniList, an open source anime database to build complimentary networks from several thousand anime entries present. This includes a production role network where edges encode how staff collaborate for the same or different titles and a small scale character network which acted as our starting point and is aimed to showcase basic insights related to them. In the character layer, the analysis is focused on the largest connected component, samples of popular series, Louvain communities, Sentiment analysis over character descriptions to name a few. In production network, we emphasized on the discrete statistics which show a heavy tailed role frequencies and collaboration counts, with high degree hubs being roles such as production assistant, key animator, episode director.

Analyzing character and production networks side by side allows us to ask whether communities and hubs in the story world align with those in creative labor and whether common characters and production roles act as structural bridges that connect otherwise separate region in anime landscape. These questions motivate the Results and Discussion sections, where the focus is on how character communities and role clusters interact and what this reveals about the joint organization of story telling and production industry in current anime.

Significance

Anime's have been airing from a long time and they build dense fictional worlds where characters and creators of those characters jointly unfold and shape the story. In this paper, we see how a production network complemented by a character (1)network, both of which constructed from AniList metadata, study how story world and industry jointly shape the anime field. By linking graph structures to text descriptions and production roles, we see how character communities, crossover figures and studio collaborations affect the anime landscape beyond the genre labels. By leveraging this dual perspective, we aim to offer a way to analyze complex media ecosystems and help explain how artists and studio choices together shape the social structure of anime worlds.

Author affiliations: [DTU]Danmarks Tekniske Universitet

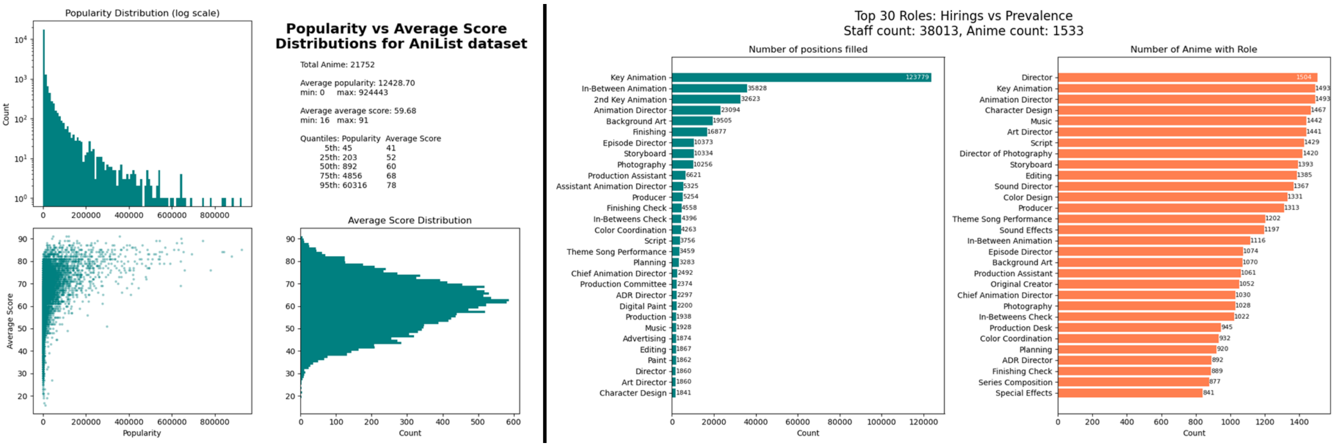

Fig. 1. Network statistics on Popularity vs Average user score (left) and Staff role distributions (right)

Results

Global properties of the network.

Role frequencies and division of labor. It has been found that the number of positions assigned per role is exponentially distributed (Figure 1 (right)). By far the most numerous positions are concerned with the production of the drawings themselves, the number of positions decreasing as the amount of work decreases as well. Key Animations are tasked with the most work, having to transform a rough Storyboard into nearly finished drawings, dictating the timing, positions and detail of key moments, which the 2nd Key Animators administer corrections to and the In-Betweeners interpolate movement between. Background Art require less work since backgrounds are often static for many frames and are moved around in post production with zooming, panning etc.

When we look at how common is it for an anime to have a specific role we gain a new perspective on production. Ranked by prevalence, roles that are responsible for coordinating others get highlighted, mainly Directors of specific departments. However we still see some roles who have individual tasks, these roles are highlighted here because their work is fundamentally necessary, such as drawing keyframes.

Role mobility characteristics. When constructing a network between roles, where the edge weights are defined by the number of staff having experience in both positions, we gain insight into staff mobility. We can observe in Figure 2 a hierarchy where Key and 2nd Key Animators are much more likely to also have experience in Animation Direction, while In-Betweeners are more commonly related to Key and 2nd Key Animation, we may be able to conclude that there are common carrier paths that staff may go through in their work (Figure 4).

When we extract the network backbone, only unusually strong connections remain (Figure 3 This helps a lot in counteracting roles that are overly numerous just based on the amount of work needed in their area like Key Animation. Now our connections should describe genuine connections between disciplines, such as License Managers and Advertising Assistance, Background Art being strongly linked to CG Set Modelling, CG Prop Modelling and Digital In-Between animation, clearly indicative of the modern trend of replacing static background paintings with computer generated scenes which can be animated in a much more cost effective way.

Discussion

Production roles as backbone of anime worlds. Our analysis of production roles suggests that anime worlds are held together not by a narrow elite of directors or creators, but by a small set of ubiquitous technical functions. The extreme skew we observed in role frequencies along with the heavy-tailed degree distributions, shows that tasks such as key animation, in-between animation, finishing, and production assistance are credited more often and collaborate with far more partners than most creative specialist roles. To us, this implies that large-scale anime production relies on a flexible pool of technical staff who repeatedly connect otherwise separate projects, effectively creating a shared production backbone that spans studios, genres, and formats

Division of labor and invisible hubs. When we compared total credits(when we say credits, we mean for work they did) against counts of unique staff per role, we found a separation between repetitive labor and meaningful leadership. We saw that roles like director, animation director, character design, music, and art director attract many different individuals who each receive relatively few credits, while a smaller core of staff cycles repeatedly through technical positions such as key animation and in-between animation. From a labor perspective, this means the people who collaborate most frequently across titles are often not those with the most authority, but rather those performing the routine tasks required for every episode. Our role-order graphs for "Director" and "Storyboard" reinforce this as both connect to a wide variety of neighboring roles, but the edge weights are moderate, pointing to broad but shallow collaboration compared to the dense clusters we found around animation work.

Familiarity networks and coordination structure. The disparity-filtered backbone (it is a method for finding most significant connections in a weighed network) provides further evidence that production coordination is built around recurring technical partnerships rather

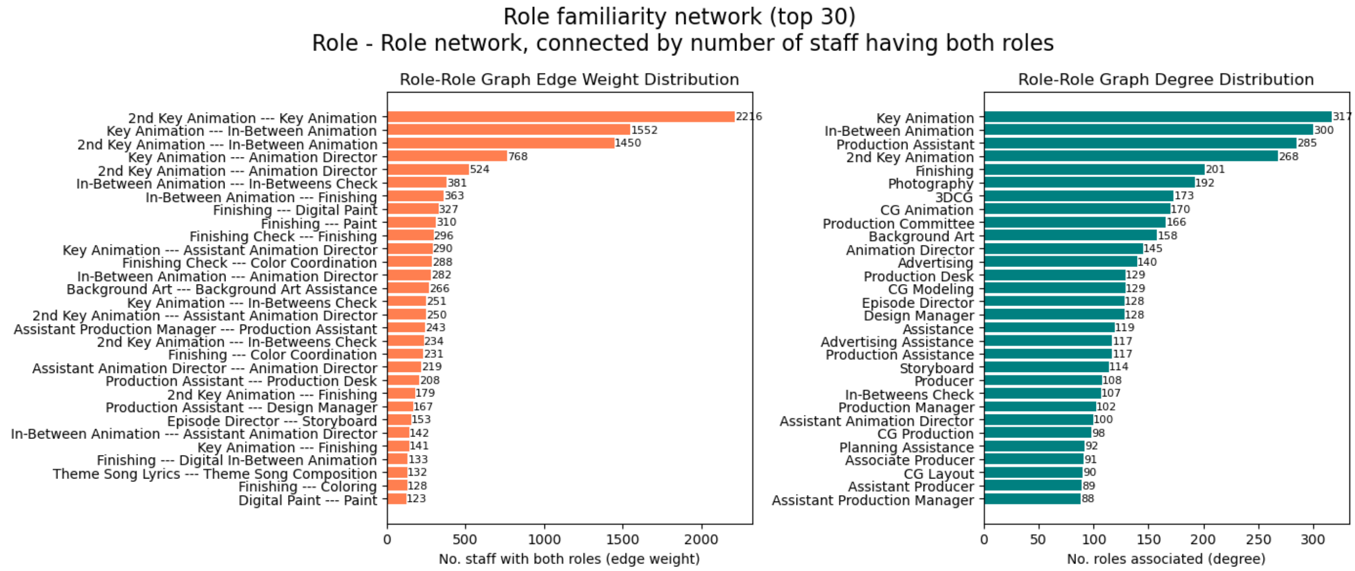

Fig. 2. Role Familiarity Network

than one-off creative pairings. After we filtered out weak edges, the top weight links almost exclusively joined combinations of animation-adjacent roles. For example, in-between animation with theme-song arrangement or finishing check with background-art assistance while the top degree roles included background art, animation director, episode director, production assistant, and production desk. These patterns suggest to us that "familiarity networks" form around roles that must repeatedly synchronize detailed work across shots and episodes, making them critical for keeping production pipelines running smoothly even when the visible creative direction changes from show to show.

Limitations and future steps. We acknowledge several limitations in our current analysis that stem from data and modeling choices. First, AniList credits are incomplete and inconsistently formatted across studios and eras, which likely under counts certain roles and may bias collaboration patterns toward more recent or better-documented series. Second, our role standardization and projection steps flatten differences between positions (like assistant versus chief) and ignore temporal order, meaning say a story boarder and animator are linked, they worked on same show but that does not mean they followed each other's instructions. Third, for computational reasons, our character-level analyses relied on the largest connected component and sampled subgraphs, which may under represent highly fragmented series or minor characters. In future work, we could add episode timings, studio names and more detailed text analysis to see how production teams shape genre or popularity. We could also expand our network model by adding layers for studios and franchises.

Methods

The notebooks containing the final code and part of the data (full data could not be uploaded on GitHub due to size constraints) used are present in *Group_69_SGAI_Project_Assignment_B*. Note, the final code and the visualizations were refined with the help of an coding assistant.

Materials and Methods

Data collection:. (2)Anime data was collected using public AniList GraphQL API using a custom Python scraper notebook we created that goes through multiple pages over media, character and staff data while respecting AniList's rate limit of 30 requests per minute. Per request you can get 50 anime or 25 staff or 25 character pages. At the current rate limiting it took about 200 hours to download the full dataset, so being able to stop, save, start the scraper was imperative. The production network was constructed from the staff entries. Information such as role (such as Key Animation, Producer, etc), specific episodes, or miscellaneous tags which enable us to reconstruct who was responsible for what during production. It is important to note that not all anime are documented equally well, we found that filtering for the number of staff listed over 100 gave a good balance for still representing shows of all popularity while giving us the data needed to reconstruct the production environment. This filtering cut down the number of anime considered for production analysis from 21752 to 1533, containing information about the 38013 staff involved and 1578 roles.

Network Construction:. Two main graphs were constructed in Python using NetworkX. The character network is an undirected weighed graph created using just a very small subset of anime (100) but has around 3900 nodes and 127k edges. All analyses were restricted to largest connected component and when necessary, to random samples of nodes defined by minimum degree cutoffs so the computations remain feasible. The production network has been chosen to be a role-role undirected weighted network, where we connect staff roles based on the number of staff that has experience in both roles. This network gives us insight into staff mobility between roles, the stronger the connection between two roles, the more staff are likely to be flexible between them.

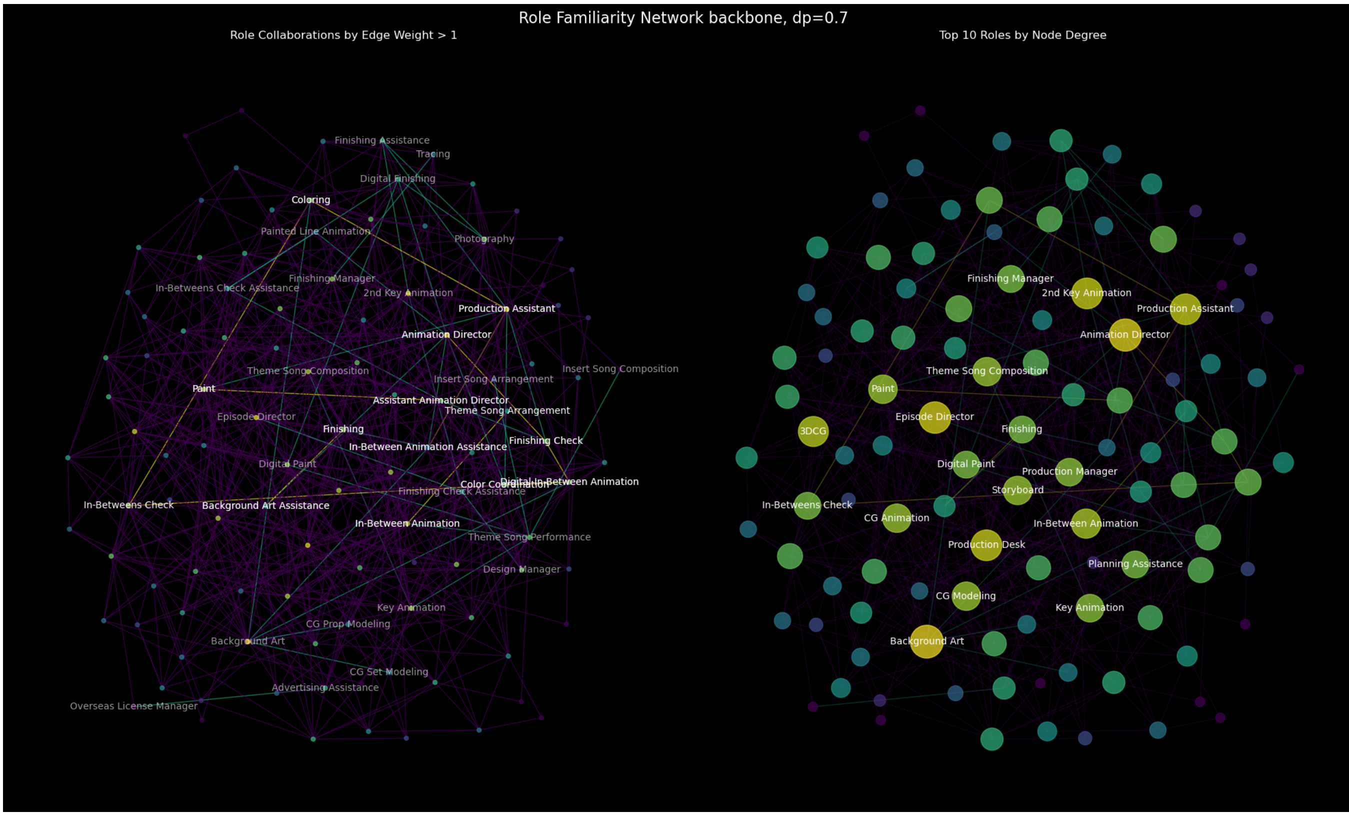

Fig. 3. Role Familiarity Network after backbone extraction

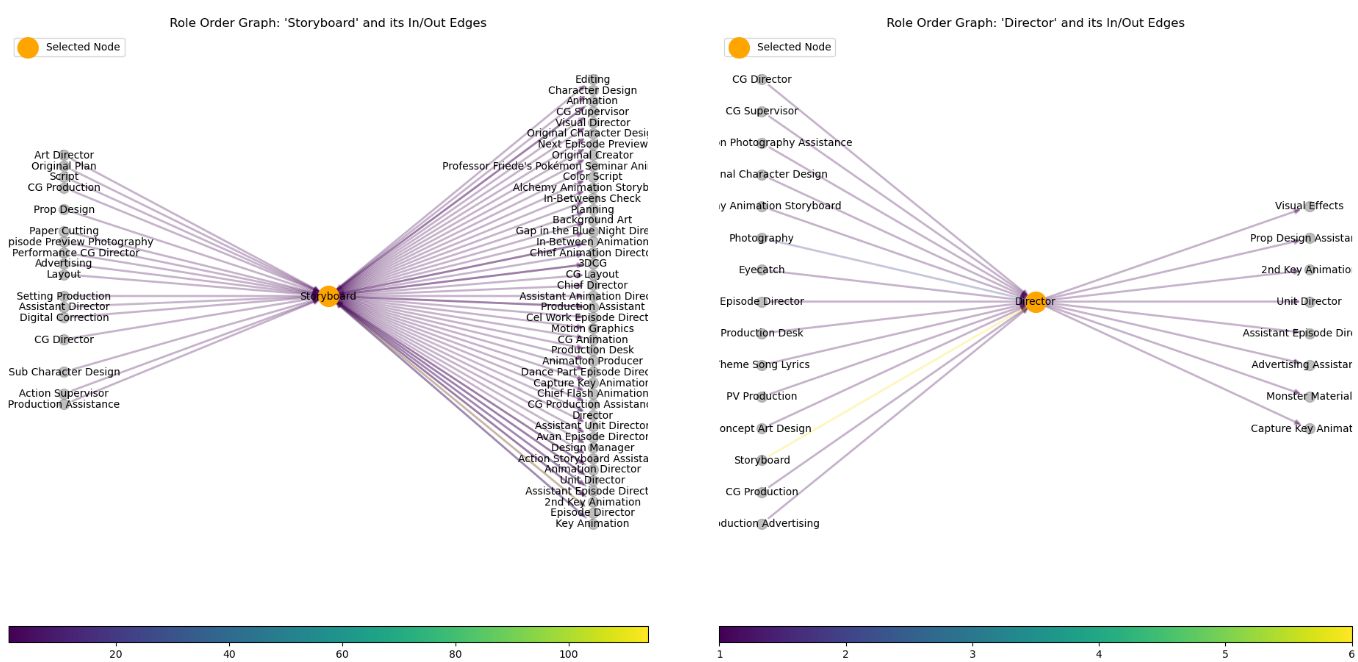

Fig. 4. Temporal staff role linking shown for two connected nodes. Edge direction based on which role did the staff get experience in first

Network and Community Analysis:. The notebook computes node and edge counts, average degree, degree distribution, size of largest connected component and global clustering coefficient. (3)Communities on largest connected component using Louvain algorithm. The script reports the number of communities, their size distribution and visualizes a 500 node induced subgraph (less cluttering, better visualization) colored by community to indicate structural groupings.

Text and Sentiment Analysis:. (4)Character descriptions are processed in python using a simple processing code. We discard very short descriptions and control characters before passing to **nltk.sentiment** which returns a compounded polarity score for every character. The scores are stored in a dict and summarized in histograms

Statistical Analysis and Visualization:. Descriptive statistics for anime popularity and average score (mean, minimum, maximum and quantiles) are computed and popularity score relationships are inspected with scatter plots as shown in Figures. All network visualizations are produced with a spring layout or force directed layouts in Network X. Figures, histograms and bar charts are generated with matplotlib and seaborn.

Table 1. Contribution table

Section	Bence	Gokul
Introduction	40%	60%
Data collection and cleaning	50%	50%
Network analysis	70%	30%
Community and Sentiment	40%	60%
Results and Discussion	50%	50%
Formatting and Visualizations	60%	40%

Declaration of Generative AI use

Generative AI (like ChatGPT) was used in editing the code, formatting the visualizations, text refinement and doing research required in order to successfully complete writing this paper. All the research decisions, analysis to perform and conclusions to derive were done by the authors.

1. MEJ Newman, *Networks: An Introduction.* (Oxford University Press), 2nd edition, (2018).
2. AniList, Anilist graphql api (https://anilist.co) (2024) Accessed: December 2025.
3. VD Blondel, JL Guillaume, R Lambiotte, E Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
4. S Loria, Textblob: Simplified text processing (https://textblob.readthedocs.io) (2023) Version 0.17.