How Do YouTubers Collaborate? A Preliminary Analysis of YouTubers' Collaboration Networks

Hiba Siraj[1], Arnav Machavarapu[2], Jiwoo Hwang[3], Kaavya Radhakrishnan[4], Scarlett Adams[4], Jinyi Kim[1], and Myeong Lee[1]

[1] George Mason University, Fairfax VA, USA
[2] Westwood High School, Austin TX, USA
[3] Thomas Jefferson High School for Science and Technology, Alexandria VA, USA
[4] Langley High School, McLean VA, USA

Abstract. Online videos such as those streamed through YouTube are largely produced by individual users rather than traditional mass media, partly due to the incentive structure of the platforms. As part of the strategy to increase the audience, many content creators collaborate with other creators to attract subscribers and diversify their content. This behavior can be conceptualized as "coopetition" as they cooperate for their channels' success while competing with one another for the limited pool of audience. In this project, we analyze data about beauty and gaming YouTubers to understand their collaboration types and network structures. The network analysis suggests that (1) the coopetition networks of YouTubers may show a scale-freeness in their topological structure and (2) beauty YouTubers cooperate with non-beauty channels more compared to gaming YouTubers, implying that YouTubers' cooperation networks may present a different level of heterogeneity depending on topics. The results inform the mechanisms of online video producers' cooperation and competition processes from an ecological perspective.

Keywords: YouTube · Network analysis · Collaboration

1 Introduction

YouTube is an online video-sharing platform that gained high popularity in recent years, with over 2.6 billion active users. YouTube offers a participatory culture that allows users to create, learn, and collaborate with each other [2]. Their videos are largely contributed by individual users [6], partly due to the incentive structure of the platforms. As part of the strategy to increase the audience, many content creators collaborate with others to attract subscribers and diversify their contents. This type of behavior can be conceptualized as "*coopetition*" [17], because YouTubers collaborate with each other for their channels' success while also competing for the limited pool of the audience. Coopetition is defined as "a strategic and dynamic process between multiple actors to create value through cooperation while competing to capture the value" [1]. Coopetition has been studied in various contexts such as business [13, 8], financial markets [3], education [12, 11], and government [10].

When it comes to online video-sharing social media such as YouTube, actors' coopetition process becomes unclear, because (1) the coordination cost to create a collaborative video tends to be lower compared to the well-studied contexts such as business and government (e.g., [10, 13]), (2) individual-level friendships may drive the collaboration heavily, rather than organizational partnership does, (3) the flexibility and fluidity of collaboration are high compared to those of formal organizations, and (4) the boundary of resources (i.e., audience distribution across topics) is blurry on YouTube, compared to industries and markets. As a growing number of individuals rely on content creation markets for their jobs, this study is critical not only in advancing our understanding of the coopetition process on social media, but also in providing implications for social media designers in sustaining and supporting their users.

To date, coopetition is relatively less studied in the context of video-sharing social media. One of the most relevant studies examined YouTubers' collaboration networks at scale by automatically detecting collaborative videos using face detection algorithms and deep neural network (DNN) models [9]. This study reported that, across the topics, YouTubers' collaboration significantly increased the number of viewers and subscribers at maximum of a 100% growth. Such collaboration networks of YouTubers were visualized as an interactive interface by focusing on the featured channels and their subscriber counts [5]. Other than these two cases, YouTube-based studies have been largely focused on the networks of friendships, subscriptions, and information flow on YouTube, with their theoretical discussions on the diffusion of information, homophily, and social contagion (e.g., [14, 4]).

While the large-scale study of YouTubers' collaboration networks and their visualization provides in-depth insights into the computational methods and network characteristics, there are still theoretical and methodological gaps. Theoretically, prior studies are still limited in providing the understanding of the coopetition processes among YouTubers by focusing on "growth" only. Methodologically, the accuracy of the network predictions (e.g., whether a video is a collaboration-based or not) is questionable as the detection of multiple people in a video does not necessarily mean a collaboration. Also, the types of networks (e.g., whether a collaboration was planned or by coincidence) are not considered, because computational methods have not yet been developed to predict the fine resolution of collaboration types. These limitations might stem from trade-offs between the content resolution and the scalabiltiy of networks.

Filling these gaps, we aim to study the coopetition processes on YouTube by examining YouTubers' collaboration networks. To understand the coopetition process, it is necessary to identify network types (e.g., regular, coincidental, or one-time collaboration) and collaborators' characteristics (e.g., non-YouTubers or not), because these features characterize the entire networks at the ecological level. Without detailed classifications of nodes and edges in collaboration networks, understanding the ecological features such as competition and legitimation would be at best difficult and, at worst, unreliable. As the initial step to exploring this problem space, we start with a small-scale collaboration network analysis by focusing on two topic categories: beauty and gaming.

2 Methods

We designed four stages to collect and process the YouTube data: collecting data using YouTube API, cleaning data using the Google Vertex AI, classifying collaboration videos, and visualizing the networks for analysis using Gephi. To investigate individual YouTubers' coopetition processes, we filtered out organizational and mass media channels and only focused on YouTube channels managed by individuals. To contrast network characteristics based on YouTubers' activities, audience demographics, and frequencies, we sampled data for the beauty and gaming categories as the initial targets.

2.1 Data Collection and Filtering

To sample videos, we selected search keywords relevant to the topics of beauty and gaming. The gaming keywords were selected based on the top 10 games (e.g., Call of Duty, Valorant, Minecraft) and the most popular gaming platforms (e.g., Roblox). Similarly, for beauty, the search keywords were chosen based on the top-10 beauty brands in the market (e.g., Channel, NARS, L'Oréal). These keywords were used in the queries to the YouTube APIs to sample the videos. Search queries were written in Python to make API requests for each keyword. The data returned included attributes such as video ID, video title, video description, channel ID, published date, and audio language. We collected 7871 videos for gaming and 6611 for beauty after removing duplicates. We also collected channel data separately using the channel IDs captured from the sampled videos.

The collected videos still included irrelevant videos that did not fall under the categories of beauty and gaming. To automate the filtering process, we sampled 100 random videos from each topic and manually tagged whether each video was relevant to the target topic or not (i.e., we tagged *yes* if the video was relevant to the topic, and *no* otherwise). The 100 tagged videos from each topic were used to train the machine learning model to classify the rest of the videos. For building the models, we used Google Cloud's Vertex AI for text classification, which selects the best machine learning model among their candidates (e.g., Gradient Boost Decision Tree [7]). The models were trained and tested using the video titles ($F1 = 0.96$) and predicted the relevancy and irrelevancy scores for the rest of the videos. We chose the videos that had higher relevancy score over irrelevancy score. A random sanity check showed the validity of the approach.

2.2 Classifying Collaboration Videos

To construct the networks, we randomly chose the seed channels. Among the pool of relevant videos, 100 videos were sampled, and their channels were listed.

Among the listed channels, we manually filtered out channels that had collaboration-based videos, and selected five channels that had largest volumes of videos. Using the five channels as the seed channels, we chose the most recent five videos in each channel. Then, we tagged whether each video was based on collaboration or not. When a collaboration was identified in a video, a link was created between the focus channel ID and the guest channel ID. Both channels were logged as nodes, and then we continued to examine the five videos from the guest channel. In this way, the manual tagging was conducted recursively as the networks grew from the seed channels, until we reach either external guests (i.e., those who do not have YouTube channels) or no collaborators.

The collaboration between two YouTubers were categorized into four network types: *one-time* for an one-time collaboration between the host and guest, *regular* for collaborations that regularly happened between the host and guest, *coincident* for a collaboration that happened without planning, and *external* for a collaboration with guests who are not YouTubers. We also tagged whether the guest channel was from the same topic as the host channel's or not.

2.3 Network Visualization and Analysis

Gephi, a network analysis and visualization tool, was used to plot the collaboration networks by modeling the channels as nodes, with their subscriber count determining the size of the node (i.e., the higher the subscriber count, the larger the node). Two heterogeneous, weighted, and undirected networks were created where both nodes and edges had multiple types, and edge weight reflected the number of collaborations between the two nodes. We color-coded the edges to show the types of collaboration between two channels, and a separate color-coding on the nodes to identify the channels from different topics. Table 1 presents the descriptive statistics of the networks.

Table 1. Descriptive network statistics.

Network statistic	Gaming	Beauty
# of Nodes	171	102
# of Edges	210	132
Average Degree of Nodes	2.46	2.34
Average Weighted Degree	4.62	3.79
Graph Density	0.01	0.02
Average Path Length	3.896	3.947

3 Results

We present the visualization of YouTubers' collaboration networks for gaming and beauty, respectively, in Fig. 1 and Fig. 2. Because the network sizes are

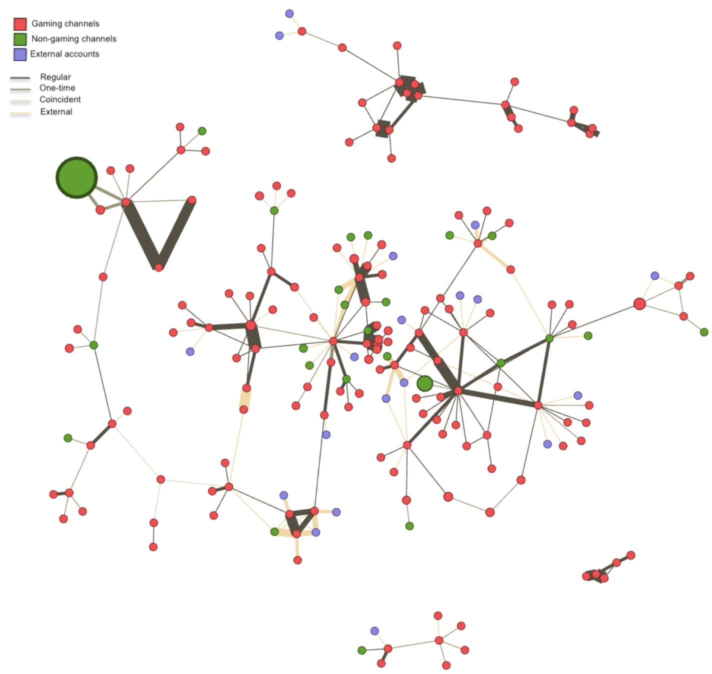

Fig. 1. Collaboration network visualization of Gaming channels.

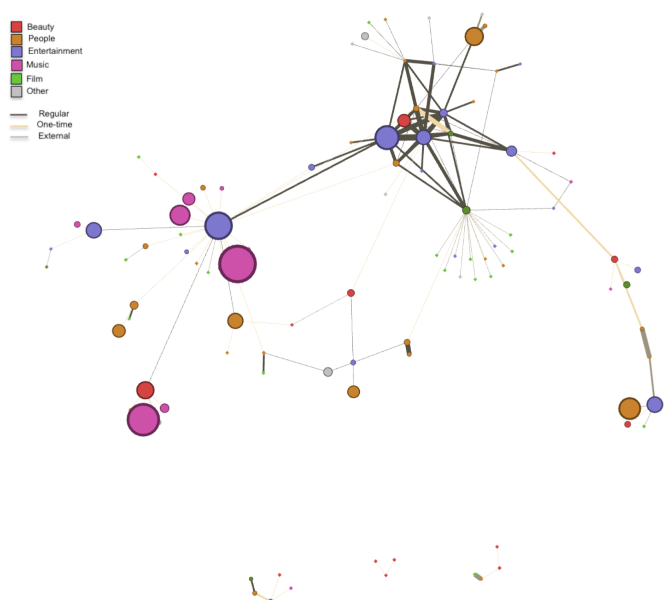

Fig. 2. Collaboration network visualization of Beauty channels.

slightly different between the two topics, it is difficult to directly compare their network densities and the entire topological structures in our samples. However, the random sampling of the nodes allows to compare the network heterogeneity and some topological features.

First, the networks show the different levels of heterogeneity between the two topics. While the gaming channels mostly collaborated with channels under the same category (color-coded in red in Fig. 1), beauty channels collaborated more with non-beauty channels. Regarding the collaboration types, the gaming channels exhibit a greater number of regular collaborations (65% of total collaborations within topic) than beauty channels do (Table 2). Also, gaming YouTubers had collaborated with external people (e.g., influences from Instagram) more than beauty YouTubers had. Meanwhile, one-time collaborations were more frequent within the beauty topic (35%) when compared to gaming (6%).

Table 2. Distribution of collaboration types.

Collaboration type	Gaming	Beauty
Regular	65.71%	48.48%
One-time	6.67%	34.85%
External	23.33%	16.67%
Coincident	4.29%	-

Second, we plotted the distribution of degree centrality of the nodes in each topic. We found that that the degree distributions of both gaming and beauty collaboration networks followed the power law distribution as shown in Fig. 3, just like other social networks. The difference is that the gaming network is more skewed towards left than the beauty network, with a potential of a higher exponent. This indicates that a fewer percentage of YouTubers in gaming may play a key role as a hub in connecting other gamers, compared to beauty YouTubers.

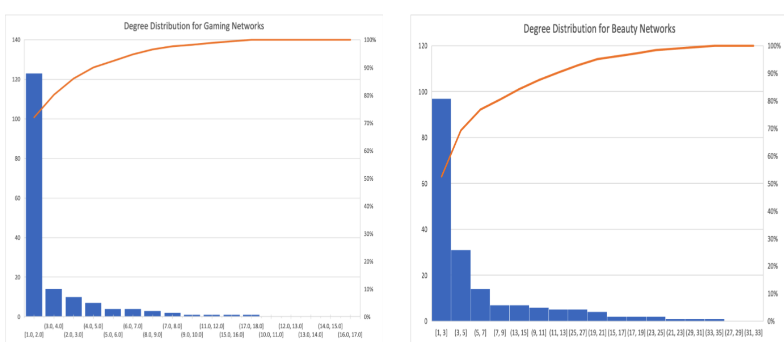

Fig. 3. Degree distribution of gaming and beauty channels' collaboration network.

4 Discussion and Future Work

Our preliminary analysis shows that YouTubers' collaboration networks tend to be scale-free, with a few channels having a higher number of collaborations compared to the rest of the network, but with the different skewness across the topics. Also, gaming and beauty topics displayed different levels of heterogeneity in collaboration, indicating that videos may present very different types of collaboration patterns across different topics. Gaming YouTubers tend to collaborate with others within the same topic, while beauty channels had more diverse types of collaborators. These findings suggest that YouTubers in different topics may have to manage the tension between cooperation and competition differently, due to varying competition structures and sizes of the audience pool. Especially, the crispness of the boundaries between ecological niches may vary significantly depending on topics, which highlight the importance of niche overlaps in understanding the competition structures on YouTube [15].

Variations in the number of external collaborators outside of YouTube provide interesting implications for the coopetition processes. While our study does not include any analyses of competition structures, YouTubers and external collaborators would minimally compete for similar resources (i.e., audience). In that case, YouTubers might not need to manage the tensions too much from a competition perspective; instead, understanding external collaborators' motivation to work with YouTubers might become an interesting topic. In our samples, the number of external collaborators was larger for gaming YouTubers, which indicates that their intrinsic motivation to play games with YouTubers might be one of the important factors in understanding the ecological processes.

As this study is in the early stage of preliminary analysis based on a small sample of beauty and gaming videos, future work will have to scale up the study by focusing on larger samples, broader topics, longitudinal aspects of the network evolution, and further numerical analyses. Expanding the sample size and time span will allow to examine (1) small-world characteristics [16], (2) ecological niches and their boundary characteristics, (3) competition structures of the niches, and (4) their effects on performances over time. Identifying these characteristics will make it possible to understand how YouTubers manage the tension between collaboration and competition, how the tension management strategies lead to their success in the context of video-sharing social media, and how video topics and network structures moderate the impact of the managerial factors. Practically, future work based on this study will benefit video streamers in planning their collaboration strategies and platform designers in supporting their users. Also, on-boarding streamers and media organizations can benefit from understanding the organizing mechanisms in planning social media strategies.

Acknowledgment

This project has been conducted partly through the Aspiring Scientists Summer Internship Program (ASSIP) at George Mason University, which aims to promote

STEM research opportunities for high school and undergraduate students. We thank the ASSIP and Dr. Andrea Cobb for their support for this study.

References

1. Bouncken, R.B., Gast, J., Kraus, S.: Coopetition: a systematic review, synthesis, and future research directions. Review of Managerial Science **9**(3), 577–601 (2015)
2. Chau, C.: Youtube as a participatory culture. New Directions for Youth Development **2010**(128), 65–74 (2010)
3. Gnyawali, D.R., Park, B.J.: Co-opetition between giants: Collaboration with competitors for technological innovation. Research Policy **40**(5), 650–663 (2011)
4. Goel, S., Anderson, A., Hofman, J., Watts, D.J.: The structural virality of online diffusion. Management Science **62**(1), 180–196 (2016)
5. Gugel, B.: Youtube universum. die vernetzung der youtuber visualisiert. https://tinyurl.com/bddhatbn (2015)
6. Hosseinmardi, H., Ghasemian, A., Clauset, A., Mobius, M., Rothschild, D.M., Watts, D.J.: Examining the consumption of radical content on youtube. Proceedings of the National Academy of Sciences **118**(32) (2021)
7. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y.: LightGBM: A highly efficient gradient boosting decision tree. In: Advances in Neural Information Processing Systems. vol. 30 (2017)
8. Kim, S., Kim, N., Pae, J.H., Yip, L.: Cooperate "and" compete: Coopetition strategy in retailer-supplier relationships. Journal of Business and Industrial Marketing **28**(4), 263–275 (2013)
9. Koch, C., Lode, M., Stohr, D., Rizk, A., Stienmetz, R.: Collaborations on youtube: From unsupervised detection to the impact on video and channel popularity. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) **14**(4), 1–23 (2018)
10. Luo, Y.: A coopetition perspective of mnc-host government relations. Journal of International Management **10**(4), 431–451 (2004)
11. Muijs, D., Chapman, C., Ainscow, M., West, M.: Networking and collaboration in education. Springer (2011)
12. Muijs, D., Rumyantseva, N.: Coopetition in education: Collaborating in a competitive environment. Journal of Educational Change **15**, 1–18 (2014)
13. Ritala, P., Hurmelinna-Laukkanen, P.: What's in it for me? creating and appropriating value in innovation-related coopetition. Technovation **29**(12), 819–828 (2009)
14. Susarla, A., Oh, J.H., Tan, Y.: Social networks and the diffusion of user-generated content: Evidence from youtube. Information Systems Research **23**(1), 23–41 (2012)
15. Wang, X., Butler, B.S., Ren, Y.: The impact of membership overlap on growth: An ecological competition view of online groups. Organization Science **24**(2), 414–431 (2013)
16. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world'networks. Nature **393**(6684), 440–442 (1998)
17. Yoo, E., Gu, B., Rabinovich, E.: Competition and coopetition among social media content. In: Hawaii International Conference on System Sciences. pp. 4338–4347. IEEE (2019)