Understanding popularity of online music videos: from a perspective of network structures

Yutian Zhao

Australian National University

u6489809@anu.edu.au

Supervised by:

Lexing Xie

Australian National University

lexing.xie@anu.edu.au

***Abstract*—Nowadays, recommendation systems are widely used in online platforms. In this work, we try to understand the effect of recommendation systems on the popularity of online music videos by performing network analysis. We start by studying the global characteristics of the network. We find the long-tailed, unequal distribution of views and significant homophily. We then demonstrate significant correlations between popularity and network structures like indegree, clustering coefficient, average neighbor indegree, and diversity of sources. We further analyze the relation of estimated network contribution and network structure. We find that videos from the same artists contribute to videos with similar popularity while videos from different artists tend to contribute to videos with higher popularity. Finally, we find videos with small and dense network characteristics benefit most from the network although the further growth of popularity may be limited. The findings of our work help to have a better understanding of the effect of the recommendation system on popularity.**

I. INTRODUCTION

In recent years, recommendation systems are more and more popular and have been widely used in online platforms or applications. Personalized recommendations are generated for different users. This has shown to be of great value and effectiveness, e.g., content providers can locate their ideal target more accurately and users can reach the content which they are interested in with less effort. Encouraged by this benefit, service providers including online audio and video streaming platforms (e.g., YouTube, Netflix, and Spotify), electronic commerce platforms (e.g., Amazon), online social media (e.g., Facebook and Twitter), and reading platforms (e.g., Microsoft news) have developed and keep updating their own recommendation algorithms to improve the quality of their service.

While the recommendation system is becoming more and more intelligent and effective, the effect of it has become nonnegligible. In daily use, a large amount of user attention is driven by the recommended content, viewing from one to the next seamlessly. This affects the popularity of the content, as the content more frequently recommended will receive more attention than those less recommended. In economics, it is found that the recommendation changes the distribution of demand from popular products to less popular ones or vice versa [1]. Considering the effect on demand, attention, and popularity, which are major interests of content providers or producers, it is worth studying what kind of effects the recommendation system has on them and how it exerts such an effect.

In this work, we will put our focus on the music videos, where the recommend system plays an important role. In fact, the recommendation system is one of the two dominant factors driving the increase of video views on YouTube, along with the active video search [2]. Our main research object is the popularity of the videos, which we choose to use the view count to measure considering it is a classic and direct measure of popularity used in many studies [1][2][3]. We aim to find the characteristics of popular and less popular videos under the recommendation system. Furthermore, what kind of videos benefits from the recommendation system. Specially, we study these questions by examining the recommendation network formed naturally by the recommendation system, which gives an illustrative representation of how videos are recommended to others and how views flow on the link. We first study the overall characteristics of the network by examining different network measurements and specific network structures like the largest connected components and bridges. Then, we relate network measurements with the popularity of videos, including indegree centrality, clustering coefficients, average nearest neighbor indegree, and diversity of sources (assortative mixing). We find they have a significant correlation with the number of views. Finally, we use the network contribution estimation from [3] to analyze the relation between network contribution and network structures. We find videos with high clustering coefficients recommended by videos from the same artists or genres benefit most from the recommendation system although the total views might be limited.

The research questions we aim to answer are:

RQ1: What are the characteristics of the video recommendation network?

RQ2: What network measurements correlate to the popularity of video and how?

RQ3: How the recommendation contribution relates to network structures?

We perform network analysis, statistical analysis, and data visualization to tackle the above questions.

The results of this work give a better understanding of how the recommendation system behaves in practice. The work gives possible reasons why a video is popular or not from the network structure perspective. And by analyzing the recommendation network, we can identify popular and potentially popular videos by their network structures. The results provide a potential way to identify the need to adjust the recommendation system and how it should be done by looking at the network structures. For example, network structures are useful to identify fairness problems, which is quite common in recommendation systems. Or one could modify the network structures to make a video benefit more from the network.

II. Related work

While most studies put attention on the accuracy of the recommendation systems, there are some studying the effect of it, many of them are in the field of electronic commerce. Hinz et al. [4] showed search and recommendation systems either lead to additional sales or substitution of sales. Oestreicher-Singer and Sundararajan [5] found the recommendation system significantly flattens the demand and revenue distributions, thus reduces inequality in demand across products. In their following paper [6], they stated that newer and more popular products 'use' the recommendation network more efficiently, and the diversity of sources amplifies the effect of the recommendation system. However, the findings from different literature might be different. Anderson [7] found that recommendation could lead to demand being shifted toward blockbusters and away from niche products because there is more chance that a recommendation link is formed for commonly co-purchased products. What is more, the results across different fields can also be different. For online social networks, Su et al. [8] pointed out that the users tend to respond positively to popular users. And for music or videos, popular bias is one of the major problems in recommendation systems, meaning that popular items are recommended more frequently than less popular ones, reinforcing the "richer get richer" phenomenon [9]. This is not only for users but also for artists [10]. Besides, there are researchers studying the diffusion of the effect of the recommendation system. They found the diffusion usually is shallow and is within few hops away from the source while there are occasional large busts of propagation [11] [12].

As seen from above, it is hard to draw a single conclusion for the effect of the recommendation system. And it is even harder to infer causal relations based on observational data [13]. This work contributes evidence using network analysis on the music video network. Network analysis is widely used in social science, for example, relationship networks on social media and citation networks in academics. Lindenlaub and Prummer [14] found that men have higher degrees and women have higher clustering coefficients in social networks, which causes men to outperform women in uncertain work environments. Hausmann and Hidalgo [15] examined the network of exports from countries and found the negative relationship between the diversity of the diversification of a country and the average ubiquity of its exports. These studies are conducted from the economics perspective and rarely relate to the recommendation system.

One similar work is done by Leem and Chun [16]. They performed correlation and regression analysis on network centrality measures and demands on the book co-purchase network. However, the experiment is purely statistical while our work uses plots as illustrative evidence.

Compared to the above work, our work tries to explain the popularity of music video under the effect of recommendation system by performing network analysis. Little work has been done on the music videos using such methods. We hope our work can give a better understanding of the effect of the recommendation system on popularity.

III. Dataset and network construction

A. *Vevo Music Graph dataset*

To construct a recommendation network of music videos, we use the Vevo Music Graph dataset curated by Siqi et al. [3] in 2018. The dataset contains 60470 official Vevo music videos, which are verified videos with a "VEVO" watermark. These videos are from 4435 artists from six English-speaking countries (United States, United Kingdom, Canada, Australia, New Zealand, and Ireland). The dataset also includes metadata useful for our research: title, channel, view count series, and its recommendation relations with other videos. One thing to notice is the data is scraped in 63 days between Sep 1, 2018 and Nov 2, 2018.

Vevo music videos are quite representative, its brand is well-known and attracts huge attention, which can be seen from the fact that 94 of all-time top 100 most viewed videos on YouTube are music, and 64 of which are distributed via Vevo [3][17]. What's more, the recommendation system YouTube used is one of the largest scale and most sophisticated industrial recommendation systems in existence, involving deep neural networks and other unknown techniques [18]. Using such an existing representative observational dataset is one way left to study YouTube's recommendation system and the popularity of its videos.

B. *Constructing Recommendation network*

There are two types of recommendation lists in the dataset. One contains recommended videos from the right-hand panel of videos that YouTube displays on its interface. The other consists of relevant videos from the YouTube Data API [3]. We follow the same setting as the original paper, using the latter list with a cutoff at 15.

To construct a recommendation network, we add videos as nodes, and a directed edge from the source video to the target video is added if the target video is in the recommendation list of the source video. There are 63 snapshots in the dataset and we constructed 63 networks. Without explicitly saying, our measurements are calculated based on the average values over 63 networks.

IV. Overall statistics of video network

A. *Metrics*

We start by computing the basic statistics of the networks. The metrics we used are as follows.

1) Indegree

Indegree is the number of incoming edges for one node, thus it is also the number of videos that recommend this video. The higher the indegree, the more frequently the video is recommended.

One metric related to indegree is called indegree correlation, which is the average indegree of incoming nodes to a node of indegree k, calculated by:

TABLE I. SUMMARY STATISTICS

	Min	Max	Mean	SD	Correlation with views
Indegree	0	870.62	5.98	20.83	0.54
Clustering coefficient	0	0.96	0.15	0.11	-0.09
Avg. Neighbor indegree	0	61.64	2.25	3.04	0.22
Same artist percentage	0	1	0.54	0.38	-0.07
Same genre percentage	0	1	0.54	0.39	0.06
Number of videos			60740		
Average number of edges			362965		

The first 5 rows are all calculated on average values over 63 days. All correlation scores are Pearson's *r* with *p* lower than 0.001.

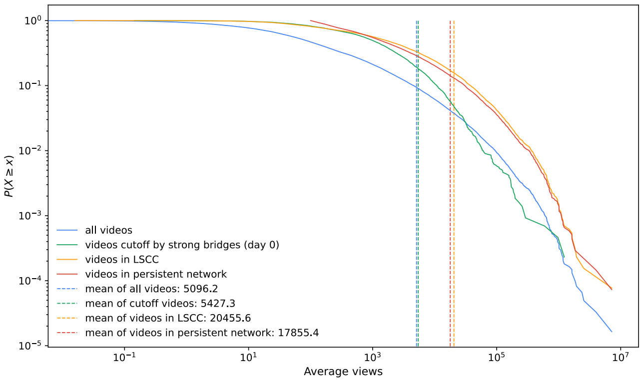

Figure 1. Average view distribution.

$$k^{nn}(k) = \sum_{k'=0}^{\infty} k' p(k'|k)$$

The indegree-indegree correlation coefficient is the Pearson *r* for indegrees of both ends. The above two metrics are to check if high indegree nodes are connected to other high indegree nodes. Outdegree is not useful in our research because we use a cutoff of 15, thus it is bounded. Another similar metric we calculated is the average neighbor indegree, the difference from indegree correlation is that this is the average indegree of incoming nodes calculated for each node.

2) Clustering coefficient

The clustering coefficient estimates how tightly the neighbors of one node are connected. The formula for the directed graph is:

$$C_i = \frac{E_i}{k_i(k_i - 1)}$$

E_i is the set of edges between neighbors of the node and k_i is the degree of the node. C_i is bounded between 0 and 1, where 1 means the neighbor nodes are completely connected with each other and 0 means completely disconnected.

3) Same artist/genre indegree percentage

As told by name, this is the percentage of indegrees from the same artist or the same genre by the total indegrees. We use this to measure the diversity of sources.

B. Long-tailed distribution

Table 1 shows the basic statistics of the video network. It is a large complex network with high derivation of indegrees.

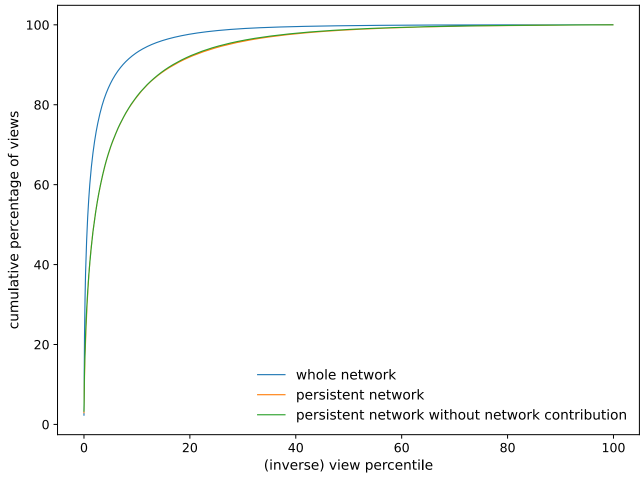

Figure 2. Cumulative percentage of views vs. (inverse) view percentile.

We further plot the complementary cumulative density function of average views in Figure 1. The so-called long tail is found, where few videos have high average views while most videos stay at the tail part of the distribution. This is shown more clearly in Figure 2 where we plot inverse view percentile (10% represents top 10% most viewed videos) with respect to the cumulative percentage of views. It turns out the top 10% most viewed videos occupy 93.1% views. The distribution is extremely unequal with Gini coefficient to be 0.946.

C. Indegree correlation

Indegree correlation is presented in Figure 3. While the average indegree of incoming nodes positively increases with the indegree, it is much fewer than the indegree, which is also true for outliers. The shows that in music video network, the videos with high indegree are acting like authority, pointed by many other videos with fewer indegrees. This result matches the average neighbor indegree in Table 1, which has a low mean value. The possible reason could be that while nodes with low indegrees are pointed by nodes with low indegrees, nodes with high indegrees are also pointed by nodes with relatively low indegrees and the value is averaged on more incoming nodes.

D. Diversity

As shown in Table 1, the percentage of incoming videos from the same artist is similar to the one from the same genre.

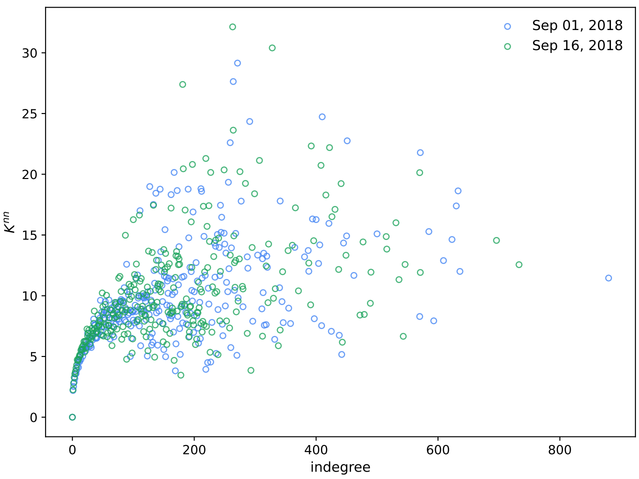

Figure 3. Indegree-indegree correlation.

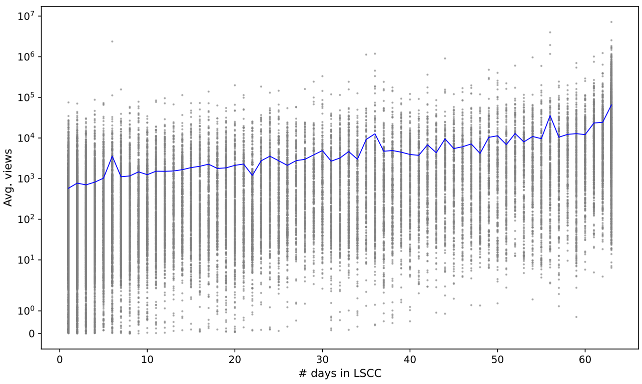

Figure 4. Popularity of nodes in LSCC. The grey dots show the distributions of average views respect to the number of days. The blue line depicts the mean.

On average, 54% of the indegrees are from the same artist/genre. We further calculate the percentage of edges 1) to the same artist, the same genre, 2) to the same artist, the different genre, 3) to the different artist, the same genre, 4) to the different artist, the different genre. The results are 1) 0.67, 2) 0.15, 3) 0.14, 4) 0.05. There are over 80% edges to the same artist or genre, showing the significant homophily of the network.

V. Relation Between Network Structure and Popularity

A. Connectivity

We first examine the global network structure. It is expected that a popular video should have an important position in the network, thus it is less likely to be disconnected from the network. Over 95% of the network are weakly connected. Instead, we compute the largest significantly connected components (LSCC). The size of LSCC ranges from 49% to 56% over 63 days. From Figure 4, we can see that average views grow positively with the number of days in LSCC.

Furthermore, we conduct a direct attack on the network by removing one strong edge in LSCC each time to see what kind of nodes being cut off using the algorithm in [19]. An edge is defined to be a strong bridge if its removal increases the number of strongly connected components. Figure 1 depicts that the videos in LSCC have higher views than the

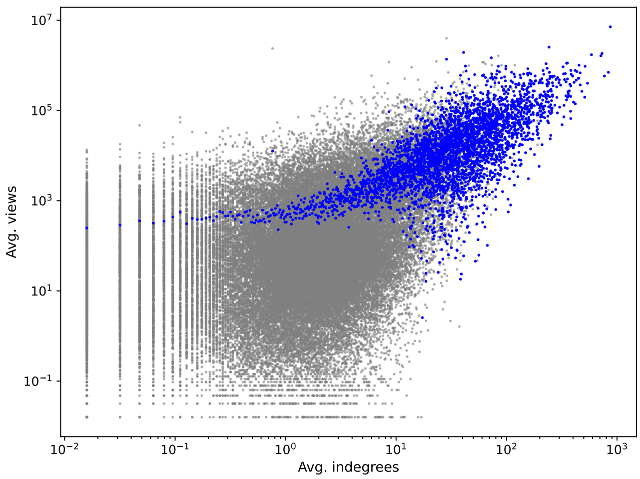

Figure 5. Average views vs. average indegrees. The blue dots depict the mean for each indegree.

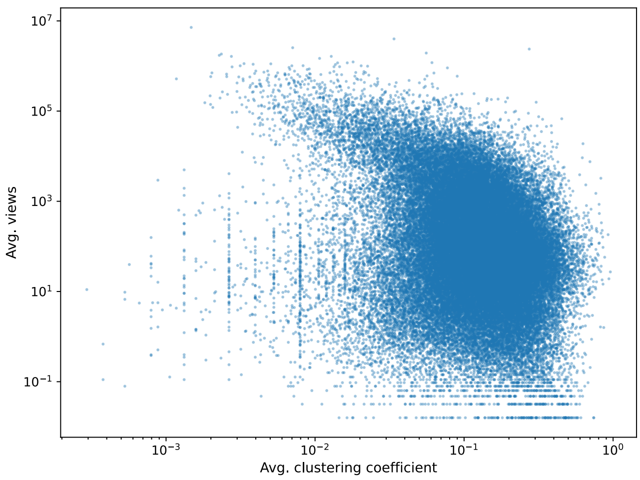

Figure 6. Average views vs. average clustering coefficient

videos in the whole network, where the nodes being cut off have similar views with the videos in the whole network. This means that videos of relatively low views tend to be cut off, thus they have lower connectivity than popular videos.

B. Indegree

Next, we examine the relationship between indegrees and popularity. The correlation test result in Table 1 shows that there is a strong positive correlation between indegrees, and views and the coefficient is 0.54. We then look at the distribution by plotting the scatter plot Figure 5. In general, views increase with indegree and as indegree increases, there is a constant increase of the lower bound of views. In other words, the variance of views is decreasing with indegrees. However, a higher view count does not necessarily mean the indegree is high. We take a closer look at the videos with high views but low indegrees. The following empirical reasons why these videos have such characteristics are found: 1) They are popular videos recommended by videos which do not belong to Vevo. The artist may have few works, or the non-official videos are more popular to be recommended. 2) The video points to many videos by the same artist acting like a hub but few of them are pointing back.

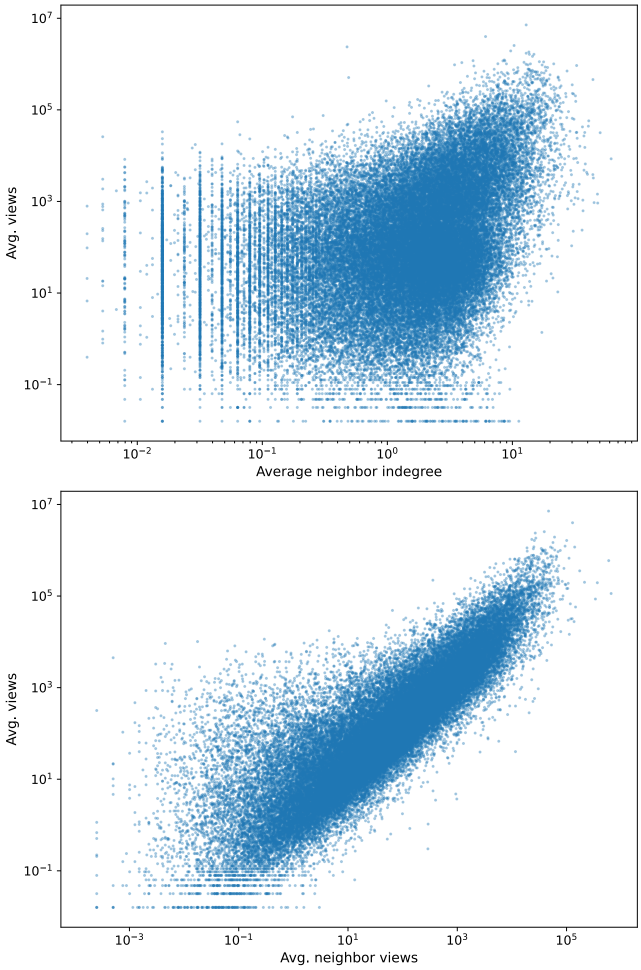

Figure 7. (Above) Average views vs. average indegree of neighbors. (Below) Average views vs. average views of neighbors.

C. Clustering coefficient

The correlation between clustering coefficients and views is opposite to the correlation between indegree and views, with coefficients to be -0.09, as shown in Figure 6. We take examples of very high clustering coefficients but with low or high views. The possible cause for a high view is the video is connected to similar hot hits and the tight connection of its neighbors ensures the amount of view flows to it, while a low view is caused by the special types of the video, for example, audios, interviews, and trailers. Although the tight connection to hot hits ensures the views, it is found for high clustering coefficients there are rarely videos with high views. This relates to our explanation of why the correlation is negative: when the neighbors are closely connected, the flow of views is 'trapped' in those videos, and the contribution from those videos are shared together, thus high clustering coefficients may limit the further growth of views.

D. Average nearest–neighbor indegree

From Figure 7, we can see that the views are positively correlated with average neighbor indegree. This correlation becomes even stronger after we consider the average views of incoming neighbors. This demonstrates that the connected videos share similar characteristics and popular videos tend to connect with other popular videos. We further examine the type of incoming edges. We divide videos into 4 groups equally by view percentile. From Figure 8, it is found that in each group the most income edges are from the same group, meaning that videos are mainly recommended among themselves. This may lead to the difficulty moving between groups and reinforce the popularity bias.

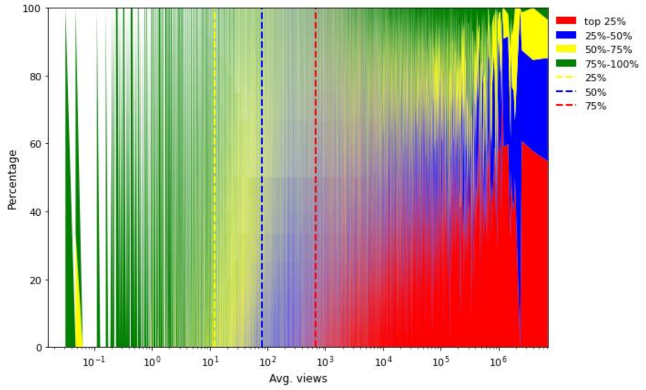

Figure 8. Percentage of incoming links from top 25%, 50%, 75%, 100% videos vs. average views. The yellow, blue, and red vertical lines represent 25%, 50%, and 75% percentile of average views.

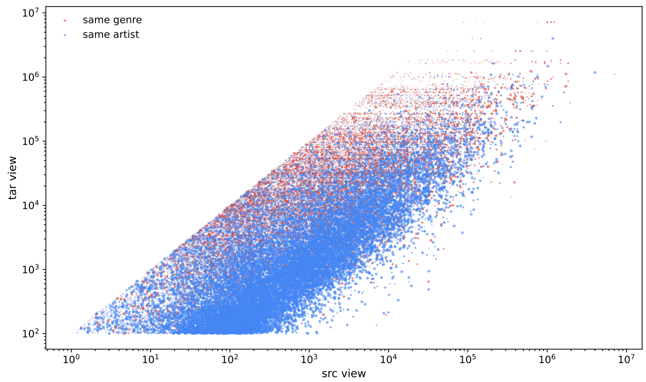

Figure 9. Edge contributions vs. source video views and target video views in persistent network. The size of the dot is proportional to the percentage of edge contribution by the total contribution the target video receives.

E. Same artist/genre indegree percentage

In Table 1, it is shown that while the percentage of indegrees from the same genre positively correlates to views, the same artist indegree percentage negatively correlates to views. This showcases the diversity matters when recommending. The users are more like to move to the video of the different artists but still the same genre. We will take a closer look at edge diversity taking account of network contribution in the next section.

VI. Relation between Network Contribution and Network Structures

A. Persistent network

After observing the correlation between network structures and popularity, we further related the effect of recommendation systems to network structures. In this work, the effect of the recommendation system is called network contribution, and it is measured by the product of the source video views and the edge weights estimated by AutoRegression used [3]. The detail of how the weights are estimated is detailed in the original paper, here we focus on the analysis of the results. One thing to note here is to estimate edge weight, the regression is run on the persistent network, which contains a subset of nodes (13710 target videos) and keeps frequently appearing edges. Compared with the whole network, the videos in the persistent network have higher views, shown in Figure 1.

B. *Network contribution analysis*

The network contribution from each source video is computed by the product of edge weight and view of the source video. We plot the edge distribution in Figure 9. Each point represents an edge, and the size of the dot is proportional to the percentage of the network contribution from the source video with regard to the total estimated contribution that the target video receives. The plot reveals the same finding as before: videos tend to connect with videos of similar popularity. Interestingly, the edges to the same artists contribute more than those to the different artists. And videos of the same artists with recommendation relationships usually have very similar views while videos tend to point to videos with higher views from different artists. Moreover, when the target videos' views are high, there are many small contributions from different artists and few bigger contributions from the same artists.

C. *Beneficial network characteristics*

To denote how much a video benefits from the network, we calculate the percentile change when network contribution is added. In other words, we first compute the original percentile of view. Then we subtract the network contribution of each of its incoming nodes and compute the percentile without the network. The percentile change is the original percentile minus percentile without network.

We take a closer look at the outliers whose percentile increases more than 50%. We found common network characteristics: they have a high clustering coefficient, relatively low indegrees (around 10), high average neighbor indegree, and all indegrees are from the same artist and genre. These findings show that the video benefits most from the network have a small and dense network structure. They are closely and stably connected to popular videos thus receive high network contribution. However, as discussed before, this kind of small and dense network characteristic may limit the further growth of views.

VII. Conclusion

In this work, we perform network analysis on music video networks. We find significant correlations between popularity and network structures like indegree, clustering coefficient, average neighbor indegree, and diversity of sources. We further analyze the relation of estimated network contribution and network structure. We find that videos from the same artists contribute to videos with similar popularity while videos from different artists tend to contribute to videos with higher popularity. Finally, we find the small and dense network characteristic benefits from the network most although it may limit the further growth of popularity.

A. *Limitation and future work*

As a network analysis research on observational data, although correlations between popularity and network structures are found, whether there are causal relations still cannot be answered. We provide empirical conclusions on the popularity of videos under the effect of recommendation system, to draw more powerful conclusions on how recommendation system leads to the results, actual recommendation systems may need to be implemented.

The methods we used are not limited to the video network, one can perform similar measurements on the artist network, which we did not show in our work.

Last but not least, there are known limitations of regression analysis. Estimation from AutoRegression may not be accurate. Also, working on the persistent network is quite limited, one can extend it to work on the whole network.

Acknowledgment

This work is completed under the supervision of Prof. Lexing Xie. Thanks for her instruction. Also, I appreciate Siqi Wu for the useful code he wrote before.

References

[1] Celma, Ò., & Cano, P. (2008, August). From hits to niches? or how popular artists can bias music recommendation and discovery. In Proceedings of the 2nd KDD Workshop on Large-Scale Recommender Systems and the Netflix Prize Competition (pp. 1-8).

[2] J Zhou, R., Khemmarat, S., & Gao, L. (2010, November). The impact of YouTube recommendation system on video views. In Proceedings of the 10th ACM SIGCOMM conference on Internet measurement (pp. 404-410).

[3] Wu, S., Rizoiu, M. A., & Xie, L. (2019). Estimating attention flow in online video networks. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW), 1-25.

[4] Hinz, O., & Eckert, J. (2010). The impact of search and recommendation systems on sales in electronic commerce. Business & Information Systems Engineering, 2(2), 67-77.

[5] Oestreicher-Singer, G., & Sundararajan, A. (2012). Recommendation networks and the long tail of electronic commerce. Mis quarterly, 65-83.

[6] Oestreicher-Singer, G., & Sundararajan, A. (2012). The visible hand? Demand effects of recommendation networks in electronic markets. Management science, 58(11), 1963-1981.

[7] Anderson, C. (2006). Why the future of business is selling less of more.

[8] Su, J., Sharma, A., & Goel, S. (2016, April). The effect of recommendations on network structure. In Proceedings of the 25th international conference on World Wide Web (pp. 1157-1167).

[9] Abdollahpouri, H., Burke, R., & Mobasher, B. (2017, August). Controlling popularity bias in learning-to-rank recommendation. In Proceedings of the Eleventh ACM Conference on Recommender Systems (pp. 42-46).

[10] Abdollahpouri, H., Burke, R., & Mansoury, M. (2020). Unfair Exposure of Artists in Music Recommendation. arXiv preprint arXiv:2003.11634.

[11] Oestreicher-Singer, G., & Sundararajan, A. (2012). The visible hand? Demand effects of recommendation networks in electronic markets. Management science, 58(11), 1963-1981.

[12] Leskovec, J., Singh, A., & Kleinberg, J. (2006, April). Patterns of influence in a recommendation network. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (pp. 380-389). Springer, Berlin, Heidelberg.

[13] Sharma, A., Hofman, J. M., & Watts, D. J. (2015, June). Estimating the causal impact of recommendation systems from observational data. In Proceedings of the Sixteenth ACM Conference on Economics and Computation (pp. 453-470).

[14] Lindenlaub, I., & Prummer, A. (2021). Network structure and performance. The Economic Journal, 131(634), 851-898.

[15] Hausmann, R., & Hidalgo, C. A. (2011). The network structure of economic output. Journal of Economic Growth, 16(4), 309-342.

[16] Leem, B., & Chun, H. (2014). An impact of online recommendation network on demand. Expert systems with applications, 41(4), 1723-1729.

[17] Wikipedia. 2019. List of most-viewed YouTube videos. https://en.wikipedia.org/wiki/List_of_most-viewed_YouTube_videos [Online; accessed Feb-12-2019].

[18] Covington, P., Adams, J., & Sargin, E. (2016, September). Deep neural networks for youtube recommendations. In Proceedings of the 10th ACM conference on recommender systems (pp. 191-198).

[19] Firmani, D., Georgiadis, L., Italiano, G. F., Laura, L., & Santaroni, F. (2016). Strong articulation points and strong bridges in large scale graphs. Algorithmica, 74(3), 1123-1147.