Reputation Formation in Online Social Media

Qian Tang, Bin Gu, Andrew Whinston

The University of Texas at Austin

Abstract

A key driver that motivates individuals to contribute to social media is to obtain reputation. This paper investigates the reputation and its formation in the context of YouTube, the largest online video site. Views of video lead to views of the provider's channel, and then channel views lead to subscriptions, through which the provider build up her reputation on YouTube. Subscriptions in turn influence future video and channel views. We use a multi-stage dynamic process to study online reputation formation. Our hypotheses are supported using panel data from YouTube on 11,000 videos and channels. Individuals first develop an interest in a creator's work. In the second stage, individuals become more interested in the creator and learn more about the author through online channels. In the third stage, interests in the creator manifest into subscriptions to the creator, which in turn influence video and channel viewership.

Keywords: online reputation, social media, YouTube

1. Introduction

A key driver that motivates individuals to contribute to social media is to obtain reputation (Wasko and Faraj 2005). But little is known on how individuals build reputation in social media. Reputation is often invisible to both researchers and contributors. As a result, prior studies frequently use proxies such as citations or references to infer the reputation of an individual. A main challenge of using such an approximation is that the approach equates reputation with quality of an individual's past work.

In this manuscript, we use a unique feature of online social media to separate reputation from quality of an individual's past work. We note that a number of online social media allows others to subscribe to future postings or creations of a particular contributor. Such subscriptions are driven by individuals' expectation of an individual's future work, i.e. reputation.

We recognition that reputation is built upon albeit different from a contributor's past work. We propose a three-stage process for reputation formation. In the first stage, individuals develop interests in a contributor's work. In the second stage, individuals interested in the work develop an interest in the contributor himself. Finally, the interests in contributor manifest into reputation that drives subscriptions to the contributor's future work. Using a large data set collected from a large social media, we show support for the three-stage process of reputation formation.

2. Research context

We study the reputation and its formation in the context of YouTube, the largest online video site. YouTube provides a platform for video creators to easily upload and share video clips on the Internet through websites, mobile devices, blogs, and email. It also provides users with a wide range of choices with millions of videos posted every day. Different from professional online video sites like Hulu, YouTube has a large amount of amateur video providers. These amateur providers share their videos through YouTube for a variety of reasons. Some of them may simply want to share a personalized video with friends or family. Some are emerging artists, such as musicians, dancers, or poets, who want to present their works to potential audience.

Each video on YouTube contains a link to the personal webpage of its creator. This creator's website is called a channel, which presents the personal profile of the creator, the social networks including friends and subscribers, and all the videos provided by the creator. A YouTube channel is similar to a traditional TV channel in the sense that both attempt to get attention from audience and cultivate loyalty among them. However, different from a traditional TV channel, a YouTube channel is it is interactive and allows users to comment on or subscribe to the channel. Comments provide a way for viewers to express their opinions, while subscription enables viewers to be informed about all future new videos from the creator. Therefore, the number of subscribers represents number of viewers who think highly of the creator so that they are willing to receive information about his future creations indiscriminatively. This measure represents the reputation of the creator.

YouTube provides a good opportunity for researchers to study the reputation formation process in online media due to its large scale and timely updated data about video views, channel views, and subscription. Views of video lead to views of the provider's channel, and then channel views lead to subscriptions, through which the provider build up her reputation on YouTube. Subscriptions in turn influence future video and channel views. It is important for researchers and video creators to understand the dynamic relationships between video views, channel views, and subscriptions. The research question of this paper is how video views lead to views of its provider's channel and subscriptions and how subscriptions in turn generate more future video and channel views.

3. Theory and Research hypothesis

Reputation is the underlying incentive for providers to contribute to online media world. As we analyzed above, people have many motives to share their videos online, such as entertaining others, seeking self-fulfillment or economic gains. Each of these motives requires providers to acquire their reputations to a certain degree. Existing research on online knowledge contribution have demonstrated that people contribute their knowledge when their contribution enhances their professional reputation (Wasko and Faraj, 2005). Internet and social network makes it possible to share their video quickly and widely. Factors that impact traditional knowledge sharing such as co-location (Kraut et al., 1990), demographic similarity (Pelled, 1996), and a history of prior relationship (Krackhardt, 1992) are no longer apparent in online world (Wasko and Faraj, 2005). This situation makes it more important for a video provider to build up her reputation online.

YouTube video providers build up their reputation based on the performance of the videos they share. Since there is no explicit co-location, demographic similarity, or prior relationship between a provider and its viewers, the only way for a provider to get attention from them is to provide high quality videos with interesting topic. If the viewers like the videos by a provider, they are more likely to be interested in viewing the provider's channel to check out her information and other videos. If the videos have greatly impressed viewers, making them willing to watch her future videos, they may end up being her subscribers. This leads to our first set of hypotheses:

Hypothesis 1.a: Video views have a positive impact on channel views.

Hypothesis 1.b: Video views have a positive impact on subscriptions.

YouTube provide a categorization of channels, where users can find channels that interest them. On the page of channel category, users can learn about the number of videos and viewers of channels. Once they click through, the channel page would present them with all the information about the channel, including provider's personal profile, statistic data of the channel, comments from viewers, videos provided, and subscribers. From the channel page, viewers can find all the videos posted by the provider and click through to watch it. They can also subscribe to the provider if they would like to be informed about updates from the provider. This leads to our second set of hypotheses:

Hypothesis 2.a: Channel views have a positive impact on video views.

Hypothesis 2.b: Channel views have a positive impact on subscriptions.

Providers' reputation in turn impact their video views and channel views as well. Online reputation mechanism can deter moral hazard or serve as signaling devices in online world (Dellarocas, 2005). In online video sharing, reputation mechanism enables users to distinguish between high quality videos and low quality videos. Number of subscriptions on YouTube signals the video creator's reputation, and this statistic is available to all users. The subscribers to the video creator are the first ones to learn about her new videos and are most likely to watch these videos. They are also more likely to view or comment on the provider's channel from time to time. Therefore, change in subscriptions reflects the dynamics in creator' reputation. This leads to our third set of hypotheses:

Hypothesis 3.a: Subscriptions have a positive impact on video views.

Hypothesis 3.b: Subscriptions have a positive impact on channel views.

4. Empirical model

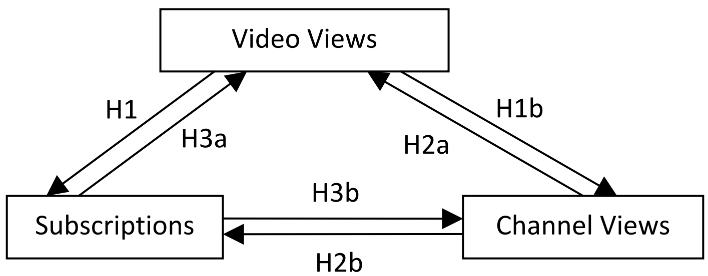

Figure 1. Research Hypotheses

Figure 1 summarizes our research hypotheses. In order to model these hypotheses, we first develop a measure of view stock to identify number of recent viewership on videos and channels. As recent views have more influence than views in the past, we give recent views more weight using the following equation to calculate view stock for videos and channels:

$$ViewStock_{it} = \theta ViewStock_{it-1} + (1-\theta)\Delta View_{it} \quad (1)$$

where we use $\theta = 0.5$. $\Delta View_{it}$ is the incremental views for video or channel i in period t.

4.1 Video Viewership

We start with the Bass Diffusion model (Bass, 1969) to study the diffusion process of video viewership. Bass Diffusion model consists of innovation, communication channels, time and the social system (Mahajan et al., 1990). All these four key elements are satisfied by our YouTube case. Based on Bass model, we use the following equation to model video viewership:

$$\Delta LgVView_{it} = \beta_0 + \beta_1 LgVViewStock_{it-1} + \beta_2 LgVViewStock_{it-1}^2 + \beta_3 LgNumOfSubscribers_{it-1} + \beta_4 LgCViewStock_{it-1} + \beta_5 LgNumOfComments_{it} + \beta_6 LgAveRating_{it} + \beta_7 LgNumOfRatings_{it} + \beta_8 LgVAge_{it} + \varepsilon_{it} \quad (2)$$

In Equation (2), $\Delta LgVView_{it}$ represents the incremental increase of viewership for video i at time t. $LgVViewStock_{it-1}$, the log of accumulative view stock for video i up to time t-1 and its quadratic form $LgVViewStock_{it-1}^2$ are adopted from Bass model (Bass, 1969). $LgNumOfSubscribers_{it-1}$ and $LgCViewStock_{it-1}$ are the log of accumulative number of subscribers and viewer stock of the provider's channel up to last period. Under hypotheses H2a and H3a, both β_3 and β_4 should be positive for subscription and channel views. The log of number of comments $LgNumOfComments_{it}$, the log of number of ratings $LgNumOfRatings_{it}$, and the log of average rating $LgAveRating_{it}$ are used to control for the impact of online word of mouth. The higher rating and the more ratings and comments a video receives, the more likely a user would like to view it (Duan et al. 2008). We also control for the age of the video using $LgVAge_{it}$. ε_{it} is the error term.

4.2 Channel Viewership

Similar to video diffusion, the channel views is also influence by accumulative channel view stock and its quadratic form (Bass, 1969). WOM also have a positive impact on channel views, which means more comments lead to more channel viewer stock. Similarly, Subscribers, who are interested in the provider, are more likely to check out update on the provider's channel. Providers' reputation is built upon the performance of their videos. If viewers like a provider' videos, they are more likely to look at her channel. The more videos and the more views providers get from their videos, the better chance they get exposure for their channels. However, even two providers with the same number of videos and the same average views for their videos could be different. Imagine one with two just so so videos and the other one with one huge success and one not so good. Which one would end up with more viewers? Since users on YouTube tend to watch what others are watching, we also propose a positive relationship between the variance of video views and channel views. Therefore, we develop the following channel diffusion model:

$$\Delta LgCView_{jt} = \alpha_0 + \alpha_1 LgCViewStock_{jt-1} + \alpha_2 LgCViewStock_{jt-1}^2 + \alpha_3 LgNumOfSubscribers_{jt-1} + \alpha_4 LgNumOfVideos_{jt} + \alpha_5 LgAveVideoViews_{jt} + \alpha_6 LgVarOfViews_{jt} + \alpha_7 LgNumOfComments_{jt} + \alpha_8 LgCAge_{jt} + \varepsilon_{jt} \quad (3)$$

In Equation (3), $\Delta LgCView_{jt}$ is the incremental increase of viewership for channel j at time t. $LgCViewStock_{jt-1}$, the log of accumulative view stock for channel j up to time t-1 and its quadratic form $LgCViewStock_{jt-1}^2$ are adopted from Bass model (Bass, 1969). $LgNumOfVideos_{jt}$, $LgAveVideoViews_{jt}$, and $LgVarOfViews_{jt}$ are the number, average views, and views variance of videos. Under hypotheses H1b and H3b, α_3, α_4, α_5, and α_6 should be positive for subscription and video views have a positive impact on channel views. $LgNumOfComments_{jt}$ is used to control for the impact of WOM. We also control for the age of the channel using $LgCAge_{jt}$. ε_{jt} is the error term.

4.3 Subscription

Based on the reputation building process, video views and channel views have a positive impact on subscriptions. We use the following equation to model subscription:

$$\begin{aligned}\Delta LgNumOfSubscribers_{kt} &= \gamma_0 + \gamma_1 LgNumOfSubscribers_{kt-1} + \gamma_2 LgCViewStock_{kt} \\ &+ \gamma_3 LgCViewStock_{kt} * LgNumOfSubscribers_{kt-1} + \gamma_4 LgNumOfVideos_{kt} \\ &+ \gamma_5 LgAveVideoViews_{kt} + \gamma_6 LgVarOfViews_{kt} + \gamma_7 LgNumOfComments_{kt} \\ &+ \gamma_8 LgCAge_{kt} + \varepsilon_{kt} \qquad (4)\end{aligned}$$

In Equation (4), $\Delta LgNumOfSubscribers_{kt}$ is the incremental increase of subscribers for provider k at time t. $LgCViewStock_{kt}$ is the log of accumulative view stock for channel j up to time t. We also include the interaction between channel views and subscription $\gamma_3 LgCViewStock_{kt} * LgNumOfSubscribers_{kt-1}$ to capture a potential complementary relationship between the two as individuals who visit a channel and see a significant number of subscribers are more likely to subscribe. $LgNumOfVideos_{jt}$, $LgAveVideoViews_{jt}$, and $LgVarOfViews_{jt}$ are the number, average views, and views variance of videos. Under hypotheses H1a and H2b, γ_1, γ_2 and γ_3, should be positive. $LgNumOfComments_{kt}$ is used to control for the impact of WOM. We also control for the age of the channel using $LgCAge_{jt}$. ε_{jt} is the error term.

5. Results

We collected panel data about 11000 videos and their providers' channels from April 1st to May 31st in 2007. Table 1 presents the results from fixed effect panel data regression using Stata 10.0. The results reveal that both number of subscribers and channel views influence the level of video views, supporting H2a and H3a. Moreover, our analysis reveals that video viewership significantly influences both channel viewership and subscriptions, supporting H1a and H1b. It also shows that channel viewership in turn influences subscriptions (H2b).

Table 1. Main results from panel data analysis

Equation	Variable	Coefficient (se.)	p-value	Conclusion
Video views	$LgNumOfSubscribers_{it-1}$	.03420(.00213)	0.000	H3a supported
	$LgCViewStock_{it-1}$	.02965(.00984)	0.001	H2a supported
Channel views	$LgNumOfSubscribers_{jt-1}$	-.03163(.00693)	0.000	H3b not supported
	$LgNumOfVideos_{jt}$	.11005(.00578)	0.000	H1b partially supported
	$LgAveVideoViews_{jt}$	.00000(.00000)	0.398	
	$LgVarOfViews_{jt}$	.00000(.00000)	0.807	
Subscription	$LgCViewStock_{kt}$	.00429(.00044)	0.000	H2b supported
	$LgCViewStock_{kt} * LgNumO$	.00039(.00009)	0.000	
	$LgNumOfVideos_{kt}$	.00958 (.00074)	0.000	H1a partially supported
	$LgAveVideoViews_{kt}$	.00000(.00000)	0.000	
	$LgVarOfViews_{kt}$	.00000(.00000)	0.016	

6. Discussions and Conclusions

The objective of this study is to take a first step to analyze reputation formation in online social media. We note that reputation formation is a multi-stage dynamic process. Reputation formation has three stages. Individuals first develop an interest in a creator's work. In the second stage, individuals become more interested in the creator and learn more about the author through online channels. In the third stage, interests in the creator manifest into subscriptions to the creator, which in turn influence video and channel viewership.

Reference

Bass, F. M. A new product growth model for consumer durables. Management Science, 1969, 15: 215-227.

Mahajan, V., Muller, E., Bass, F.M. New product diffusion models in marketing: A review and directions for research. *Journal of Marketing*, 1990, 54:1-26.

Dellarocas, C. Reputation Mechanisms. In T. Hendershott, ed., *Handbook on Economics and Infdormation Systems*, 2006.

Duan, W., Gu, B., and Whinston, A. B. Do online Reviews Matter?-An Empirical Investigation of Panel DataThe dynamics of online word-of-mouth and product sales — An empirical investigation of the movie industry. *Journal of Retailing*, 2005, 84(2): 233-242.

Krackhardt, D. The Strength of Strong Ties: The Importance of Philos in Organizations, in Organizations and Networks: Structure. Form, and Action, N. NohrJa and R. Eccles (Eds.), Harvard Business School Press, Boston, 1992: 216-239.

Kraut, R. E., Egido, C, and Galagher, J. Patterns of Contact and Communication in Scientific Research Collaboration," in Intellectual Teamwork, J. Galagher, R. E. Kraut, and C. Egido (Eds.), Lawrence Eribaum Associates Inc., Hillsdale, NJ, 1990.

Pelled, L. H. Demographic Diversity, Conflict, and Work Group Outcomes: An Intervening Process Theory. *Organization Science*, 1996, 7: 615-631.

Wasko, M. M., Faraj, S. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, 2005, 29(1):35-57.