Informational Cascades and Contagion in Online Social Networks

Jeong-ha Oh, Anjana Susarla and Yong Tan

Department of Information Systems and Operations Management,
Michael G. Foster School of Business, University of Washington, Seattle
{jhoh, asusarla,ytan}@u.washington.edu

Keywords: Informational Cacades, Digital Content, Social Networking

Motivation and Research Question

A plethora of models of content creation such as Facebook, MySpace, Second Life and YouTube are facing robust growth accelerated by Web 2.0 technologies and standards and offer a fascinating avenue for exploring the role of social influence online. For businesses seeking to monetize social search and digital content, in particular, it is increasingly important to understand how users interact and participate in these settings, and to deploy techniques of reaching out to customers such as viral marketing that are aimed at harnessing the power of the collective.

It has been long recognized that individuals make decisions not as rational and atomistic utility maximizers, but are actively influenced by the tastes and decisions of others. The impact of others' actions influencing individual decisions has been studied under a variety of phenomenon labeled variously as 'herding' (Banerjee 1993), 'social learning' (Ellison and Fudenberg 1993), and 'conformity' (Bernheim 1992). In an influential paper Bikhchandani et al. (1992) discuss how such phenomenon result from 'informational cascades' where a large number of individuals converge in their behavior, leading to large-scale popularity in a trend or idea. Informational cascades occur when (i) imperfectly informed agents, acting sequentially, choose the same action as their predecessors ignoring their private information, and (ii) individuals face a simple binary choice decision.

The You Tube model makes it very easy for content creators to upload, edit and process entertainment content online. You Tube provides tremendous opportunities for social interactions via visiting and commenting on other users' channels as well as direct observation of other users' decisions in sampling content. The user-friendly interface, network externalities from initial adopters and the growth in popularity of online video catapulted YouTube into a dominant position on the Internet with over 500% growth in 2008. Indeed, video search on YouTube accounts for a quarter of all Google search queries in the U.S in 2008. The nature of viewership in user-generated models of digital content naturally lends itself to an analysis of the formation of cascades. It has been estimated that almost 70% of Youtube videos were only viewed once, while the top 20% of videos managed to gain more than half of all views. A video is an experience good and the YouTube model of "broadcast yourself" implies that there are limited marketing or promotional efforts that provide informative signals to prospective viewers, unlike that of say, a movie, or album that is reviewed by news media or promoted through radio or other marketing channels.

Bikhchandani et al (1998) posits that 2 essential conditions to triggering a cascade: local dependencies or localized learning and a binary decision characterized by uncertainty. In other words, whether or not a video can morph into a runaway hit that results in an informational cascade depends on social influence structured through a network of connections. Cascades can be triggered when small initial shocks affect highly connected nodes in a social network (Watts 2006). In this paper we are interested in how the social network structures on You Tube influences the formation and propagation of informational cascades that lead to the phenomenal popularity of some videos. The two questions we explore are:

- What characteristics of content creators and content adopters cause online informational cascades?

- How does network position of users cause cascades to tip?

While prior literature has investigated 'word of mouth' issues in the context of online electronic commerce (Chevalier and Mayzlin 2006), the focus in this paper is to identify the process of sequential information transmission and local network effects that dictate formation of cascades. It is posited that observational learning from network neighbors may be heavily influential in subsequent cascade formation – a phenomenon termed 'localized conformity'.

The data and context

We use a data set of video information and user information collected from YouTube.com. We identify three layers of network structure on You Tube. A friend network is built upon mutual agreement between users and constitutes a unidirectional network. Subscription networks do not rely upon mutual agreement have a bidirectional network structure. The third network layer is the network among videos formed when the videos are listed as related videos. We explore the role of several network measures to assess the importance and connectedness of each node in a social network. A user that has friended or subscribed to a channel can favorite a video or post comments about a video posted by that channel. At the same time, content creators can engage in promotional efforts on their channels as well as interact with users by responding to posted comments. Finally, channels can also provide informational cues through the video description. Thus, we can observe not only the networked structures of connections but also the interactions between networked actors through exploring comments on You Tube. The ratings of videos and the video description constitute a relatively "public" signal about the quality of content while interactions between channels and viewers as well as actions such as friend "favoriting" a video are relatively private signals. Comments by early adopters and actions by channels to promote content early in the life of a video can result in dynamics that lead to cascades. Therefore the nature of networked interactions lends itself to an analysis of models of cascades that posit that individuals may ignore their public information and thereby trigger an informational cascade.

One of the empirical challenges in identification is that most of the literature deals with the impact of informational cascades on individual behavior (e.g., Brock and Durlauf 2001), while our focus is on the role of networked interactions that lead to a cascade. Another challenge is that users' actions such as commenting and favoriting videos may themselves be influenced by the networked structure of interactions rather than independent actions by users – which gives rise to the well known "reflection problem" (Manski 1993) in analyzing social interactions. The methodological contribution is in combining text-mining techniques to conduct sentiment analysis with aggregate population dynamics that result in social effects resulting from local network neighbors' choices. By analyzing the textual information provided in the video comments we can explore how informational cues about (uncertain) video quality increases the growth in views, which is especially critical early in the life of a video. We also build upon literature that has analyzed social interactions such as Manski (1993) and Bajari et al (2006) to conduct exclusion restrictions to identify how the actions of an agent are distinct from that of other agents in his social network. The prominence of an actor posting comments about a video combined with the type of informational cues provided in the comments allows us to identify the threshold conditions that lead to informational cascades. Given the rich structure of social network structures on You Tube, we can also analyze which of the social network interactions has a greater impact on the success of a video.

Contributions

While commentators have traditionally highlighted the importance of opinion leaders in accelerating the acceptance of new products and posited that marketers should target "influencers", our study suggests a different mechanism, that early signals lead to a path dependent diffusion curves, which, when combined with the fundamental uncertainty in predicting success of digital content can disproportionately influence the aggregate success of products.