

CITY UNIVERSITY OF HONG KONG

香港城市大學

Jumping over the Network Threshold:  
Information Diffusion on Information Sharing  
Websites

跨越網絡的門檻：  
信息分享類網站上的信息擴散

Submitted to  
Department of Media and Communication  
媒體傳播系  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy  
哲學博士學位

By

Wang Chengjun  
王成軍

September 2014  
二零一四年九月

## **Abstract**

As a ubiquitous phenomenon, information diffusion has long been studied in communication research. Recently, accompanied by the rise of Web 2.0, information sharing websites (ISWs) have emerged on the World Wide Web as a new platform for information diffusion. ISWs can be defined as social media through which users submit, share, and comment on news or other types of information, which underscores the transformation of information diffusion as a participatory process. Technically, ISWs are usually based on social networking services (SNSs), information aggregators, and search engines. The evolution of information and communication technologies (ICTs) and their interplay with the public have transformed the underlying mechanisms of information diffusion, which have important implications for diffusion research and call for systematic investigations.

This research starts from a puzzlement in online information diffusion —contrary to the proposition of global diffusion, researchers keep finding that large-scale diffusion is rare and fragile. From the perspective of threshold models, the present dissertation aims at systematically investigating information diffusion on ISWs. Drawing on classic diffusion theories (e.g., news diffusion theory) and communication models of information flow (e.g., the ABX model), I extended the ABXC model to the ABXCT model as a general framework to incorporate and explain the emerging ICTs for information diffusion on ISWs, networked individuals, and the interplay between them.

The rise of digital media makes the “digital fingerprints” of human communication publicly available. This, in turn, grants us an opportunity to unobtrusively investigate the underlying diffusion mechanisms. Using the digital traces of information diffusion on Sina

Weibo, Digg, and YouTube, three studies are carried out in order to reformulate the threshold hypothesis of interpersonal effects, to compare collective gatekeeping with interpersonal effects, and to test the temporality hypothesis for online information diffusion research, respectively.

First, the study of Sina Weibo reformulates the threshold hypothesis for interpersonal effects in light of both threshold models and the J-curve model of information diffusion. Using the data of information diffusion on Sina Weibo, the findings confirm the threshold hypothesis of interpersonal effects. In particular, for information of public interest, interpersonal effects have a positive influence on diffusion size; for information of personal taste, interpersonal effects have a negative impact on diffusion size. The depth of diffusion networks is limited and the temporal diffusion curves are characterized by strong bursts. Thus, information diffusion on microblogs is both structurally and temporally constrained. To make information go viral means to break the bottleneck of local clusters and to spread information to diverse communities.

Second, the study of Digg primarily focuses on collective gatekeeping. ISWs provide a testing ground for the debate on the primacy of collective gatekeeping over interpersonal effects. Emphasizing interpersonal effects on ISWs, recent studies of online information diffusion generally overlook another player — information aggregators of ISWs. Through the process of news aggregation based on collaborative filtering, ISWs carry out a “collective gatekeeping” process that differs from news disseminations by mass media and news transmissions among individuals. Using the news voting data collected from a social news website (Digg), I compare the relative effects of interpersonal sources and collective

gatekeepers on information diffusion, and find that collective gatekeeping is the primary driver for online news diffusion. Considered together, the nature of ISWs itself has a strong impact on the outcome of information diffusion. In the case of Digg, the mechanism of public voting rather than local social influence dominates the size of information diffusion. In addition, the daily diffusion curves of news diffusion on Digg are also characterized by strong bursts.

Third, to understand the origin of bursts in public attention, I shift to a more general framework by proposing the hypothesis of temporality in the study of the diffusion of YouTube videos. Public attention is the extent to which individuals collectively allocate their attentions to cultural products across space and time. In addition to fragmentation and polarization, recent studies have shown that the temporal distribution of public attention is also characterized by strong bursts. Measuring bursts using peak fraction, this study aims at systematically studying how bursts of public attention are related to the popularity of YouTube videos (both lifetime and size of diffusion), and how it grows out of system recommendation, social influence, use of search engine, and mobile communication. The results show that, first, bursts are negatively related to the popularity of online videos; second, bursts are negatively related to searches, but positively related to system recommendations and the use of mobile devices; third, compared with videos in other categories (such as entertainment videos and education videos), news videos experience stronger bursts, while copyright-protected videos exhibit weaker bursts. In all, the findings suggest the temporality underlying various influential factors results in the bursts of public attention. Viral videos tend to seek continuous, incremental improvements, not great ups and downs. Thus, “going

viral” means long-term growth rather than a temporary eruption. If the fragmentation of public attention implies “winner takes all”, the temporality of public attention implies that the winner grows steadily over time.

To summarize, the present dissertation contributes to the research of information diffusion in four aspects: first, it adopts the ABXCT model as a general theoretical framework to incorporate the emergence of ISWs, and to map concrete studies about information diffusion in a systematic way; second, it formulates and confirms the threshold hypothesis of interpersonal effects; third, it conceptualizes collective gatekeeping as an alternative driving force of information diffusion; fourth, it investigates the burst of public attention in information diffusion, and formulates the temporality hypothesis. Overall, by illustrating the driving forces, the hidden patterns, and the underlying principles of online information diffusion, the present dissertation underscores the communication power of networked individuals and new communication technologies in the age of social media, and sheds light on the puzzlement about the limited size of online information diffusion.