International Journal of Research in Marketing 34 (2017) 100–119

Contents lists available at ScienceDirect

IJRM

International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

Modeling the role of message content and influencers in social media rebroadcasting

Yuchi Zhang [a,*], Wendy W. Moe [b], David A. Schweidel [c]

[a] Leavey School of Business, Santa Clara University, 500 El Camino Real, Santa Clara, CA 95053, United States
[b] Robert H. Smith School of Business, University of Maryland, 7621 Mowatt Ln, College Park, MD 20740, United States
[c] Goizueta Business School, Emory University, 1300 Clifton Rd NE, Atlanta, GA 30322, United States

ARTICLE INFO

Article history:
First received on September 3, 2014 and was under review for 7 months
Available online 30 August 2016

Guest Editor: Steven M. Shugan

Keywords:
Social media
Social influence
Text mining
Twitter
Bayesian estimation

ABSTRACT

We develop a model that examines the role of content, content-user fit, and influence on social media rebroadcasting behavior. While previous research has studied the role of content or the role of influence in the spread of social media content separately, none has simultaneously examined both in an effort to assess the relative effects of each. Our modeling approach also accounts for a message's "fit" with users, based on the content of the message and the content of messages typically shared by users.

As an empirical application, we examine how Twitter posts originating from top business schools are subsequently rebroadcasted (or retweeted) by other users. We employ an individual-level split hazard model that accounts for variation in rebroadcasting decisions related to (1) content, (2) the content-user fit and (3) the influence of other users. We find that the rebroadcasting a message depends not only on message content but also on the message's fit with a user. Our analysis also yields measures of influence and susceptibility to influence for each user, which can be used to identify influential social media users. We demonstrate how our approach can be used to evaluate different types of seeding strategies designed to increase the reach of social media messages.

1. Introduction

In 2013, consumers posted 139 billion messages to Twitter, a popular social media messaging platform. Of these messages, about 24% were rebroadcasted (i.e., retweeted) by other Twitter users (Leetaru, Wang, Cao, Padmanabhan, & Shook, 2013). Rebroadcasting, or "sharing" in other settings, is an extremely important activity for marketers, as they attempt to disseminate marketing communications about their brands on social media. Recently, there has also been great interest in studying rebroadcasting behavior, specifically in regards to who is more likely to retweet messages on Twitter (Lambrecht, Tucker, & Wiertz, 2015; Suh, Hong, Pirolli, & Chi, 2010; Zaman, Fox, & Bradlow, 2014). By understanding the drivers of rebroadcasting activities, managers could more effectively control and disseminate communications on social media. That is, by generating greater rebroadcasting activity, managers can reach potential consumers who are not actively following the content posted by the firm. However, beyond identifying characteristics of individuals who are more likely to rebroadcast, little is known about why those 24% of tweets are subsequently rebroadcasted. Nor is it known what the factors were that affected users' decisions to rebroadcast these messages while the remaining 76% were not rebroadcast.

* Corresponding author.
E-mail addresses: yzhang6@scu.edu (Y. Zhang), wmoe@rhsmith.umd.edu (W.W. Moe), dschweidel@emory.edu (D.A. Schweidel).

http://dx.doi.org/10.1016/j.ijresmar.2016.07.003
0167-8116/

One potential explanation is that the message content was key to users' rebroadcasting decisions. A number of researchers have examined how message content affects the spread of word-of-mouth (Berger, 2011; Berger & Milkman, 2012; Berger & Schwartz, 2011; Heath, Bell, & Sternberg, 2001). They find that certain types of content are more likely to be rebroadcast than others.

Alternatively, influential users could have played a role in affecting the rebroadcasting decisions of other users. Researchers have studied the role of influence in both offline and online contexts (e.g., Watts & Dodds, 2007). As more data become available through online platforms, such work has also focused on measuring the effects of influence as a function of a user's ability to influence the behavior of others and a user's susceptibility to social influence (Aral & Walker, 2012; Trusov, Bodapati, & Bucklin, 2010). If firms can entice rebroadcasting by influential users, others may subsequently follow and amplify the reach of the original marketing communication.

While the above-mentioned streams of research have contributed significantly to our understanding of social media behavior, these streams have been developing in parallel with very little integration. Our goal in this research is to propose a model of social media rebroadcasting behavior that integrates the various factors shown to influence rebroadcasting behavior. That is, can we model both the role of message content and influence simultaneously, allowing us to assess the relative impact of each? Furthermore, an integrated model allows us to investigate the impact of content-user fit, a measure that considers the interaction between the message content and user preferences. The modeling integration of content, influence, and content-user fit allows us to potentially uncover scenarios where disseminating the most viral content, on average, may not generate the most rebroadcasting activity for the firm due to the unique preferences of the audience and influentials who follow the firm's messages. Thus, this research provides managers with a method to properly determine a social media rebroadcasting strategy that is specific to their context and audience base.

To represent rebroadcasting activity over time, we employ a split hazard model to model an individual's decision of whether or not to rebroadcast a social media message and, if so, when. In addition to incorporating content- and user-specific variables, we develop a new measure that captures the fit between content and user. The fit measure is especially important because it allows us to investigate the extent to which managers should tailor the message content for individual users in their audience. We also account for social influence in our modeling framework to capture both a user's ability to influence the behavior of others and a user's susceptibility to the influence of other users.

One challenge with modeling social media data is the unstructured nature of the text that constitutes the data. Consistent with prior research, we employ a text analysis procedure to incorporate unstructured textual data into our model. We apply Latent Dirichlet Allocation (LDA) methodology to analyze the text and identify the topics featured in the content of each message (e.g., Blei, Ng, & Jordan, 2003; Tirunillai & Tellis, 2014). We then integrate the LDA results into our proposed model structure, thereby allowing us to empirically identify the number of topics or key themes in the posted comments. The identification of key themes is important in our ability to characterize individuals' interests and their fit with posted content, which marketers can use to develop targeted messages.

Our results suggest – consistent with the literature – that the rebroadcasting of social media messages varies with the content of the message. However, more importantly, our analysis provides evidence that the fit between the message content and the audience's interest (as characterized by the content of their previous tweets) is a significant driver of rebroadcasting behavior. Our analysis reveals that rebroadcasts by influentials affect the rebroadcasting behavior of other users who are susceptible to influence. However, we observe considerable variation across users both in terms of their ability to influence other users and their susceptibility to the influence of others' rebroadcasting.

To illustrate the managerial relevance of our model, we simulate how tweets are retweeted over time by varying (1) the content of the original tweet, (2) the audience composition, and (3) the seeding strategies in which we induce different users to retweet soon after the original post. These simulations provide managers with insight into whether to focus their efforts on viral content, seeding strategies, or tailoring content to fit their audience's preferences. For example, firms could identify the appropriate content to disseminate to their unique audience base and develop a message by using words that have a high probability of being associated with that content topic (as determined by LDA). They could also identify influentials and disseminate messages with words that are associated with the topics generally broadcasted by the influentials. Our results suggest that tailored messages designed to match the preferences of the audience are most effective when there is less heterogeneity in the audience's interests. Our analysis also provides evidence that while seeding influential users can increase the rebroadcasting of a message, there may be a limit to its effectiveness because the interests of influential users may differ from those of the broader user base. Our findings highlight the importance of modeling content-user fit in a user's rebroadcasting decisions and consequently in a firm's social media marketing strategies.

The remainder of this paper proceeds as follows. In the next section, we discuss the related literature and show our conceptual framework of social media rebroadcasting. Next, we describe the data we employ in our empirical analysis and detail our approach for converting the unstructured text to quantitative metrics. In the Model section, we develop our modeling framework and describe how we incorporate user-specific differences, content effects, content-user fit and the role of social influence into a unified model. We then describe the results of our empirical analysis and conduct a series of simulations in which we evaluate the effect of alternative message design and seeding strategies on rebroadcasting activity. We conclude with a discussion of the implications of our research.

2. What drives rebroadcasting?

Social media rebroadcasting is the act of sharing content found online with social peers. Academics have a keen interest in understanding this activity, as marketers seek to leverage customers to propagate the firm's content in order to reach a larger