Abstract

Social media has attracted attention as an information channel for content generated in heterogeneous Internet services. Focusing on social media platforms, we examine the factors behind social transmission with content crossover from other services through hypertext link (URL). We investigate the effects of source influence and peer referrals on diffusion outcome and address their variations in the case of content crossover.

Using 3,504,831 diffusion processes identified from real transaction data in Twitter, we demonstrate the existence of source and peer effects on content diffusion. These effects vary when shared content is linked with an external service by a URL. Source influence and peer referral have a positive impact on cascade size in the content dissemination process. In the case of content crossover, the impact of source influence decreases. However, the impact of peer referrals increases in the process of external content dissemination. Implications for researchers and managers are discussed.

1. Introduction

With the advent of online media, consumers are no longer passive information seekers but, rather, active co-producers of value and meaning [1]. These new media present both opportunities and challenges to firms: In terms of opportunities, inter-consumer communications have more impact on customer decision processes than traditional media marketing [2]. Furthermore, online media collect data on user behavior so that firms can monitor customers’ reactions to information as it is disseminated. However, inter-consumer communications do not occur in the boundary between firms and consumers, but, instead, are coproduced in consumer networks [3].

On social media, people create an impressive amount of content, such as articles, YouTube videos, product reviews, and pictures. People also share content others make through social media in forums, blogs, and online social networks. In information diffusion among people in social media, people do two things: they construct their own networks and they use them as information channels. Content is selected by the user and “pushed” both from and to his/her personalized online social network.

Social media consists of heterogeneous Internet services and it is continuously integrated with newly emerging services. These new services have their own ways of creating, sharing, and consuming content [4]. User-generated content is not only consumed in the service in which it is created but also shared with other services by URL, the one-click path to content defined in Web 2.0 technology. In this web of content, the group of Internet services called social networking services (SNSs) plays the important role of content aggregator [5]. In this category of services, users network as defined by the service [6]. Content is broadcast through user-defined networks that are exploited by users who intend to share external content via URL. We define the term content crossover as a sharing of external content by URL on SNSs.

This research examines the interaction between factors identified by prior research on information diffusion through social media: source and peer effects versus URL. Prior works examine factors in the diffusion process of Twitter, especially source influence, that is, the impact of the user who initiates the diffusion process, and peer referrals, that is, the impact of messages that are added to content by people participating in the diffusion process. However, these quantitative empirical works overlook an important element of social media: its heterogeneous nature of diffusion in social media. Content is often delivered by a URL. There can be a difference in the impact of factors that affect the diffusion process between the process with a URL and the process without a URL.

Our study focuses on the diffusion process in social media, including content crossover into social networking services. Two research questions integrate both user-driven diffusion and the heterogeneity of
services, two elements of social media that extend previous research: (a) What factors drive the diffusion process in social media? (b) Do these factors vary in the case of content crossover from external services?

We conduct an empirical study using 233 million real transaction data generated by 1,203,196 Korean users of Twitter, a well-known social media service, and collected by Kwak et al. [7] with the snowballing method. We analyze these authors’ raw data to trace content dissemination by users according to their posting and sharing behaviors.

The results demonstrate source and peer effects on content diffusion and that these effects vary when shared content is linked from an external service by a URL. Source influence and peer referral have positive impacts on the cascade size of the content dissemination process. In the case of content crossover, the impact of source influence decreases. However, the impact of peer referrals increases in the process of external content dissemination.

The remainder of this paper is organized as follows. Section 2 introduces the background of the research with related literature. Section 3 reviews the relevant literature and builds our theory and hypotheses. Section 4 describes the data used in this study. Section 5 presents our research model and explains each variable to support the hypotheses. Section 6 discusses the results of our analysis. Section 7 discusses the implications of our findings. Finally, Section 8 concludes by providing the study’s contributions, and limitations.

2. Background

Our research argument targets the difference between internal and external content dissemination. Both processes occur in social media platform services, or SNSs, through a user-driven diffusion process called social transmission [8].

2.1 Online Social Transmission

Our paper focuses on the key user behavior sharing online content through social media. Recent research on word-of-mouth and viral marketing suggests that the outcomes of sharing behavior can be linked to beneficial results for firms, such as more sales [3, 9, 10] and better stock market performance [11]. Consumer-to-consumer interaction is generally more effective than traditional marketing activities [2].

In online media, the phenomenon of consumer interaction is quite common. For instance, a movie marketing company might create a teaser clip and upload it to YouTube. From that point on, the diffusion process is mainly driven by customers. People can share the video on their own blogs, post reviews on their Facebook profiles, share their experience about the video in an online community, or send the video’s link to members of their personal networks on Twitter. People exposed to content through various social media channels can then consume it and share it in their own social networks, extending the dissemination. The process is recursive, decentralized, and collective. The outcome, measured as view counts, or hits, varies according to not only content characteristics but also social factors that affect the social transmission process.

Although this social transmission process is difficult to observe or measure due to its endogeneity and the heterogeneity of the Internet services involved, SNSs such as Facebook and Twitter have received much attention from both practitioners and researchers. There are the several reasons to focus on SNSs to research diffusion process in social media: First, this category of services represents a significant share of Internet page views. Of all the social media services in May 2012, Facebook constituted one fifth of U.S. users’ page views [12] and Twitter was third by page view rank for U.S. users [13]. Second, online social networks in SNSs function as diffusion channels for online content created elsewhere in social media. Third, most of the services in this category have data that are valuable to marketers and researchers in tracing the path of social transmission.

2.2 Twitter

Twitter, one of the most popular Internet social media services, has emerged as a new medium. Twitter allows users to post (or “tweet”) messages on any topic with a 140-character limit. Postings, called tweets, can include hypertext to link to other Internet content. The users can both “follow” other users or be followed. A follower receives postings from any user he or she is following, much like subscribing to a newspaper. Unlike most SNS sites, such as Facebook, the relationship defined in Twitter requires no reciprocation, thereby constructing a direct network. YouTube also provides this type of networking.

Based on text and symbolic communication, user interaction to content on Twitter involves well-defined markups and usages. For example, the @ sign followed by a username identifies the user. Postings beginning with @username indicate a response to the user whose identifier is the username. The markup RT indicates a retweet. When users share another’s content, they add RT @ followed by the source’s username, as in RT @username to the original content before forwarding it to their followers. Finally, the ‘#’ sign followed by a word comprises a hashtag, which is used to denote trends. The retweet is a social transmission process in
which users spread information by choice, that is, a user-driven diffusion process, and the main focus of this paper.

2.3 Related Work

This study focuses on the diffusion process in Twitter. Extensive exploratory research has been conducted on online social networks, for examples, online information diffusion investigating social media services, such as Facebook [14, 15], YouTube [16-18], blogs [19, 20], and Twitter [21-26]. Using the extensive diffusion research background on Twitter as an empirical basis for our study, previous research topics on Twitter cover influential identification [23, 24], political communication [19, 20, 25], and information diffusion [15, 21-23, 26].

Cha et al. [24] compare different measures of user influence in Twitter and suggest using the number of followers, retweets, and mentions. Kwak et al. [23] use the number of followers and retweets, and PageRank in the follower network, with similar comparison results between user ranks. Both studies conclude that user influence measured by the number of retweets and that measured by the number of followers are different.

Different areas of social network research have conducted interviews with Twitter users to determine what gets retweeted and why. Due to heavy amount of communication data on Twitter messages, Boyd et al. [22] research various aspects of retweeting with a qualitative method.

Given the various characteristics of retweeting in previous work, we ask intuitive research questions about diffusion by retweeting on Twitter. Suh et al. [26] conducted exploratory research to determine the factors behind retweeting. They examine both content features (URLs, hashtags, referencing other users) and contextual features (the numbers of followers and followees). The results show that URLs and hashtags are significant factors, as well as the number of followers. Bakshy et al. [31] investigate the number of followers and followees, as well as the past value of the number of being retweeted, which is the dependent variable in both studies. They limit their observations to messages that contain URLs.

Most of the rich and extensive research investigating the information diffusion phenomenon in social media has been conducted from a technical perspective with an exploratory approach. Our research adopts a theoretical foundation and research variables affecting the diffusion outcome from previous works. Prior research measures source influence as the source user’s number of followers, and examines the number of followers from a technical perspective, but we build an experimental research setting to understand source influence in the diffusion process on Twitter.

The next section describes our theory and hypotheses, and how the latter are derived from related works that support our research questions.

3. Hypotheses

We examine the factors that drive content diffusion in this context and then we investigate their variation on internal and external content dissemination. Prior studies of the diffusion process have investigated such factors as the characteristics of the users who post content to services (the source effect) [18, 27, 28], the participation of other users in the process (the peer effect) [14, 29], and the characteristics of the content itself (the content effect) [25]. Our research model uses the source and peer effects as the main effects, and the content effect as the moderator. We build a theory for the main effects and discuss the content effect as a contingency in Section 3.2.

3.1 Source Influence and Peer Referrals

3.1.1. Source effect. A stream of studies on the diffusion process in the context of social media is based on the idea that a small number of users influence an exceptional number of their peers. Prior research show that influential users can influence others’ virtual product adoption [27], online activity [30], and diffusion of video content [18].

A stream of studies on the diffusion process in the context of social media is based on the idea that a small number of users influences an exceptional number of their peers. Prior research show that influential users can influence others’ virtual product adoption [27], online activity [30], and diffusion of video content [18].

The number of ties in a social network is a well-established measure of individual-level influence in that network. Both the validity of the measurement and the theory linking it to influence have been researched for word-of-mouth transmission [30] and in the context of online social networks [27, 31].

To examine network connectivity for content dissemination in social media, we use Katz’s definition of interpersonal influence [32]: (a) the personification of certain values (or “who one is”), (b) competence (“what one knows”), and (c) strategic social location (“whom one knows”). Given the way people connect with each other in SNSs, a user with a high degree of interpersonal influence can attract a high number of people.

Influential SNS users have reputation, credibility, and popularity [27, 33]. Such social capital positively
affects others’ intentions to share these users’ content. Content delivered by influential users is influenced by their knowledge and their network. A core user’s high degree of credibility and popularity leads to high levels of direct exposure for the content that user shares through the network.

H1: The network connectivity of the content source in a social media service has a positive impact on the content’s diffusion.

3.1.2. Peer effect. Content adoption and the diffusion process in social media take place through user interactions. The term peer referral refers to messages that are added to content by people participating in the diffusion process. Referral messages are shown with the original content when it is shared with others through online network services. Some social media services, such as Facebook and Twitter, support a functional feature that allows users to attach personalized referral messages when sharing online content. The referral message is shown with the content and its source.

Aral and Walker [14] investigate the effect of peer referrals on socially contagious adoption behavior in online social networks. Even though the authors consider peer referrals as viral features of the diffused product, network adoption recursively activated by the adoption of others in the personal network is a different type of social transmission. As a result, peer referrals facilitate product adoption by influencing peers to adopt and share the product.

Content sharers express their emotions about content, provide information about it, or create shared meaning through further peer referrals. Such participatory interaction can draw others into the interaction. As an alternative explanation, content that induces referrals from users who then retransmit it has greater viral potential than content without interactivity.

H2: Peer referrals of content in a social media service have a positive impact on the content’s diffusion.

3.2 Content Crossover

Social media consists of various Internet services. Each service stores content created or copied by users to their own area of the Internet. A basic technical element of Web 2.0, the hyperlink, or URL, links heterogeneous services by providing a one-click path to any content, anywhere on the Internet. The way users consume information and content on social media has been called ‘news that finds us’ [34]. In such social media services, if a user invites others into his or her online network, they push information and content to the original user. Information and content can thus be created or it can be screened by users for their subscribers. A user can also publish content for subscribers and vice versa. Content pushed by multiple users is sorted chronologically or by popularity for immediate feedback. This information aggregation is displayed via the user interface, such as the Facebook’s newsfeed and Twitter’s timeline.

Computer-mediated communication theory can be used to interpret diffusion behavior in social media. Jones et al. [35] conducted a study on user behavior on Usenet, an early type of Internet community, and find that users’ cognitive load processes information exchanged via computer-mediated communication through public cyberspace. To induce a large user response, content should be simple to understand and to consume.

This theory applies to the content-consuming environment of social media services, in which consumable messages are generated by selected users. The creation and sharing of content in social media are computer-mediated communication processes. Users of social media services communicate with each other. A message that includes a URL requires more cognitive resources for user consumption due to value uncertainty [29] and the effort required to click on the link and see where it leads. As suggested above, source and peer effects reduce this uncertainty and facilitate sharing through interactive communication.

Source influence on social transmission is based on the authorship of the internal content and the user’s reputation in the service. Users post online content as prescribed by the service, such as videos on YouTube, articles in blogs, or personal life information in Facebook.

With external content, the source effect of the reduction of uncertainty of the content decreases because external content is not often produced by the source. Except for some users, who build reputations as external content providers, source influence decreases in the case of content crossover through a URL.

H3A: The impact of the network connectivity of the content source in a social media service decreases when the content is linked by a URL to another social media service.

In the case of peer referral, added messages can reduce uncertainty and facilitate sharing since the messages are personalized to the online social network members. This effect increases in the case of content
crossover due to the greater uncertainty caused by blind content linked by a URL.

H3B: The impact of peer referrals of the content in a social media service increases when the content is linked by a URL to locations in other social media services.

4. Data Description

To examine the factors that drive content dissemination in social media and how they vary in content crossover, we need data describing user sharing in SNSs, including the source who originally posted the content. Social media services use various means of content sharing—such as clicking the Share button in Facebook, scraping in blogging services, and “retweeting” in Twitter—to facilitate online social transmission. We choose Twitter, one of the best-known microblogging services, because of its simple content and data availability. In addition, content crossover using a URL is common on Twitter [26].

We use the Twitter data collected by Kwak et al. [7] who analyzed the dissolution of online relationships on Twitter.

4.1 Retweet versus RT

Retweeting is the forwarding of another user’s tweet to one’s own network of followers, using the markup RT @ followed by the source’s username. Indicating the source user gives that user credit for the content. When people on Twitter retweet others’ tweets, they can add their own message to express their own emotions, opinions, or referrals about the content. This rule of “commenting on a retweet” is not technically enforced; rather, it is learned through the community and evolved socially, like the other markups. These cultural and functional features of Twitter encourage users to share content in a measurable manner. Comments on the process of retweeting tweets can also affect diffusion outcome as peer referrals.

When content is retweeted by sequential users, each user’s markup indicates the source of each comment, in sequence, in the content. Korean users call retweets with comments RTs, distinguishing them from retweets that have no comment.

In November 2009, Twitter added an interface that enables a one-click retweet function [36, 37]. The feature shows the original author and consolidating retweeters under the message but does not allow comments. This interface allows users to track retweets and avoid redundancy in the timeline. However, the new feature is compatible with the old-style retweet functioning with markups. Regardless of whether a comment is added, retweeting involves RT @username as the markup. Using the markup RT @username, we analyze the social transmission process from individual retweets of both types.

4.2 Data Collection and Processing

The data used in the study consist of two parts: (a) user network data generated by the behavior following another user, and (b) content data posted by each user to the timeline.

Kwak et al. [7] focus on Korean Twitter users and define the network boundaries accordingly using the snowball method of network sampling. Identifying Korean letters by character sets, the authors select only Korean users whose (a) tweets, (b) profile, (c) locations, or (d) usernames are in Korean, for a total of 1.2 million Korean-speaking users. Due to the volatility of follow network data, the authors crawled daily network snapshot data 41 times from June 25 to July 15, 2010. Regarding the content data, all tweets written by the focal users can be collected retrospectively using Twitter application programming interface. We thus start with 41 network snapshots of 1.2 million users during 9 weeks and their 233 million tweets.

We analyze the tweet data to identify chains of retweets as social transmissions. The unit of analysis is diffusion process-level. Each process has a source user who posts the original tweet and the number of users who adopt the message by retweeting defined as the diffusion outcome. The process can include comments and the original message can contain hypertext to external content on another social media service. We measure content crossover by whether the message contains a URL or not, assuming that all URLs link to the valid content in external services.

First, we carry out the selection process based on the markup RT @username. We exclude messages starting with @username because that indicates a reply to content, even if the RT markup is used in its body. A total of 9,530,917 tweets were retweeted without comments and 9,648,352 tweets were retweeted with comments.

After selecting the posts involved in social transmissions, we assemble the data to reconstruct the social transmission process. We conduct a text-based analysis using 32 characters (22.8% of the 140-character content limit), including the username and initial part of the source message, as a source message identifier. This method, based on user-added markup, does not exhaustively or accurately identify online user behavior but it demonstrates 90% sensitivity and specificity in research on message dynamics of Usenet.
We obtain 4,733,317 message chains through this message aggregation. We exclude message chains started before the first network snapshot was gathered, on June 25, 2010, because they have no available measurement of source influence (see Section 5.1). Finally, we obtain 3,504,831 message chains as a sample for further analysis.

5. Model

As discussed, we examine the factors that influence diffusion outcome in social transmission and the difference on their impact between internal and external content dissemination. The number of adopters is the common dependent variable in the diffusion of innovation research. Our study considers adoption behavior as being equivalent to diffusion behavior. Because the number of adopters is a non-negative integer value, we use Poisson regression.

5.1 Variables

A group of postings diffusing the same source message, known in prior research [23] as a retweet tree, is the unit of analysis in this study. The number of links constructing each tree is the dependent variable that measures diffusion outcome. While past research has used the number of exposed users as a measurement of diffusion outcome, we do not for several reasons: (a) This quantity is difficult to measure precisely due to the temporal variation of online social networks as information channels, (b) message delivery does not directly indicate exposure of the content delivered, and (c) the measurement that we use includes the number of users who transmit the content among the number of users exposed to the content, that is, a user who transmits content is exposed to the content but not vice versa. Thus, the number of users who transmit content is a more conservative measure of diffusion outcome.

Source influence is measured by in-degree centrality defined in the follow networks, in other words, the number of followers of a source user. In-degree centrality is the number of connections initiated by other users. In SNSs, which function as central information channels of social media, users create social networks as prescribed by the service. An edge in online social network works as an information channel through which content generated by a user is shown to other users. A user with a higher level of in-degree centrality is followed by more users and has more channels through which followers can directly react to the content.

Follow networks vary according to Twitter users’ behavior. To measure source influence on the content diffusion behavior of other users, we should use the network data at the time the user diffuses particular content. We have only 41 follow network snapshots in a sequential order. We use the source user’s in-degree centrality computed based on the most recent network snapshot among the 41 at the time the first diffusion was made, as an approximation. Because there is a high degree of dispersion in the number of followers among source users in the sample diffusion process, we use a log-transformed value of the number of followers to control for skewness.

Twitter users can add a comment in front of their markup, according to socially defined rules. Peer referral is defined as a dummy variable that equals one if at least one comment is added by a peer in the diffusion process and zero otherwise.

To distinguish crossover diffusion processes from other processes that convey internal content, we use a dummy variable that equals one if the source message contains hypertext to external content, and zero otherwise. The source message is defined as a part of tweet right after the latest RT @username markup. The source message can be modified during the diffusion process (e.g., broken telephone effect [22]). We define the variable only by the source message of the earliest retweeting in each diffusion process to avoid the problem.

The definitions and names of the variables used are summarized as follows:

- **DOUT**: the number of users who retweet the message
- **SRC_INF**: the number of users following the source user of the process at the time of the initial retweet
- **PEER_REF**: a dummy variable indicating whether or not at least one comment is included in the process
- **URL**: a dummy variable indicating whether or not a URL was included in the source message of the earliest retweeting in the process

The summary statistics for all variables are shown in Table 1.

Table 1. Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOUT</td>
<td>2.31</td>
<td>12.56</td>
<td>1</td>
<td>3394</td>
</tr>
<tr>
<td>SRC_INF</td>
<td>2068.37</td>
<td>8,687.70</td>
<td>1</td>
<td>284,222</td>
</tr>
<tr>
<td>logSRC_INF</td>
<td>5.38</td>
<td>2.31</td>
<td>0.69</td>
<td>12.56</td>
</tr>
<tr>
<td>PEER_REF</td>
<td>0.77</td>
<td>0.42</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>URL</td>
<td>0.35</td>
<td>0.48</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Observations: 3,504,831
5.2 Analysis

We use a Poisson regression model due to the discrete nature of the dependent variable. Stieglitz and Dang-Xuan [25] use a Poisson regression model in their study of political communication through Twitter. They focus on retweet behavior and use the number of times a message is retweeted as the dependent variable. Because this study has the same focus and dependent measurement, we adopt the same regression model. We regress diffusion outcomes over the source, peer effects, and interactions in the case of crossover diffusion:

$log(DOUT) = \beta_0 + \beta_1 \log(SRC\_INF) + \beta_2 PEER\_REF + \beta_3 URL + \beta_4 \log(SRC\_INF) * URL + \beta_5 PEER\_REF * URL + \varepsilon$

6. Results

We conclude that source influence and peer referrals exert a significant impact on the number of adopters in online social transmission. We then show that these effects vary in the context of external content dissemination with hypertext. The summary results with four hypotheses and relevant coefficients from the Poisson regression model are shown in Table 2. The detailed results of the Poisson regression model are presented in Table 3.

Table 2. Summary of Results

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Relevant Coefficients</th>
<th>Result Summary</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal Contents Dissemination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H1: Source influence</td>
<td>log(SRC_INF)</td>
<td>Positive: 0.1924 (0.1992)</td>
<td>Yes</td>
</tr>
<tr>
<td>H2: Peer-referral effect</td>
<td>PEER_REF</td>
<td>Positive: 0.3399 (0.5017)</td>
<td></td>
</tr>
<tr>
<td>External Contents Dissemination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H3A: Decreasing Source influence</td>
<td>log(SRC_INF) x URL</td>
<td>Negative: -0.0116</td>
<td></td>
</tr>
<tr>
<td>H3B: Increasing Peer-referral effect</td>
<td>PEER_REF x URL</td>
<td>Positive: 0.3895</td>
<td></td>
</tr>
</tbody>
</table>

Note that the numbers in parentheses indicate coefficients of model without interaction terms.

The first row of Table 3 provides evidence that source influence, measured by the number of followers, has a significant impact on the diffusion outcome as measured by the number of adopters. Content that is introduced or created by a user of higher in-degree centrality is more likely to be diffused by more users through social transmission on Twitter service (H1 is supported).

Hypothesis H2 predicts that the content with peer referrals is more likely to be diffused by more users than without them. The second row of Table 3 supports this argument with the significant variable PEER-REF (H2 is supported).

The third row of Table 3 shows that content with a URL has less dissemination than content without URL.

To compare internal and external content dissemination, we examine the interaction effects between crossover and each factor presented in H1 and H2. The regression coefficient for these interaction terms is significant, as shown in the fourth and fifth rows of Table 3. For H3A, we test whether the impact of source influence on cascade size for social transmission decreases in the case of external content dissemination. In this case the regression coefficient for H3A, the interaction term between source influence and content crossover, should have a negative sign. The result is confirmed by the results in the fourth row of Table 3 (H3A is supported).

The fifth row of Table 3 show that the sign of the interaction term between peer referrals and content crossover is positive. This validates H3B, indicating that peer referral has higher impact on diffusion outcome in content crossover than in internal content dissemination on Twitter.

7. Implications

Our study presents empirical support to identify factors affecting the content dissemination process in social media. The results also contrast the underlying factors in the social transmission of internal and external content. Prior research on online
dissemination addresses the influence of source users [10, 18, 27, 38] and peer effects such as consumer opinion [9], adoption behavior [29], and referrals through the personalized online social network [14]. This study provides an integrated perspective and empirical support to determine the factors of online user-driven diffusion process on Twitter.

When firms wish to disseminate content, they should consider source influence and peer referrals to reach the desired extent of diffusion. Especially in the case of content, the effect of source influence on the dissemination process decreases but the effect of peer referrals increases. Firms should understand this at the time of content creation. For example, they may insert in their content more sympathetic elements for target users in the case of videos stored in other services such as YouTube because these elements can draw more users' referrals through the diffusion process.

From the information systems perspective, social media services should consider these characteristics in existing services. New services can announce themselves by diffusing their content through existing services. They may encourage influential users in existing services to use or request preview of content in the target services to reduce uncertainty in content selection.

8. Conclusion

Given the growing popularity of social media, particularly SNSs with online social networks as information channels, we first consider online social transmission as a user-driven diffusion process. Based on social factors in the diffusion process, we derive source and peer effects on the social transmission process. Second, we address the heterogeneous nature of Internet services in social media. We examine how factors in the diffusion process vary in the case of external content with URLs. With these research objectives, we conduct empirical analysis using a large amount of Twitter transaction data.

The results show that source and peer effects impact content diffusion, and vary when shared content is linked by URLs to external services. Source influence and peer-referrals have a positive impact on cascade size in the content dissemination. In the case of contents crossover the impact of source influence decreases. However, the impact of peer-referrals increases in external content dissemination. Our findings indicate that firms that wish to diffuse information through social media or enter the social media with new services to provide new ways of creating and sharing content should understand the nature of the social transmission process.

Our study contributes to three areas of research. First, we provide empirical support for existing diffusion theory based on communication theory as applied to user behavior in social media. Second, it contributes to the empirical literature on online consumer behavior and presents managerial implications for practitioners using social media as an information channel. Third, regarding the issue of using a large amount of data, this study is an early example of exploiting extensive real transaction data collected from computer science literature.

As with any empirical work, the external validity of the results is considerably limited. The social transmission process on Twitter investigated here is a part of the general diffusion process in social media. Although the factors we examine are valid in other environments, provided by other services in social media services, the diffusion outcome may differ due to contingencies in these services. In addition, we did not explicitly examine content characteristics as important factors. Since source and peer effects have direct impact on diffusion outcomes in the social transmission process, these effects may be mediated by the content effect. Influential users may screen better content through their own insight and content that draws peer referrals may be more viral than content without peer referrals. This more detailed model with a mediator is a possible focus of future research.

For future research, we propose an extension of this study with more general contexts focusing on content characteristics. Content discrimination may be a good starting point to address what channel they use to disseminate their content, including determining the content-type that is proper for their objective.

9. References


