

The Impact of Message Characteristics on Online Viral Diffusion in Online Social Media Services : The Case of Twitter

Youngwoo Nam
Korea University Business School
Korea University
(owoo@korea.ac.kr)

Insoo Son
Korea University Business School
Korea University
(insoo114@korea.ac.kr)

Dongwon Lee
Korea University Business School
Korea University
(mistlee@korea.ac.kr)

.....

In this paper, we explore the information diffusion mechanism under social network environments by investigating the effect of message characteristics on the volume and speed of retweeting in Twitter, a popular online social media service. To this end, we select eight main keywords (i.e., ‘무상급식’, ‘반값등록금’, ‘나가수’, ‘평창’, ‘김연아’, ‘박태환’, ‘아이폰’, ‘갤럭시’) that have been popular on online social media in recent days. Each keyword represents various social aspects of Korea that recently grab people’s attention such as political issues, entertainment, sports celebrities, and the latest digital products, and eventually holds distinctive message characteristics. Analyzing the frequency and velocity of retweeting for each keyword, we find that more than half of the sample messages posted on Twitter contain personal opinions for the certain keyword, but we also find that the tweets which include objective messages with hyperlink are the fastest ones when being retweeted by other followers. In overall, when being retweeted, the group of messages related to the certain keyword present distinctive diffusion patterns and speed according to message characteristics. From academic perspective, the findings in the study broaden our theoretical knowledge of information diffusion mechanism over online social media. For practitioners, the results also provide managerial implications regarding how to strategically utilize online social media for marketing communications with customers.

.....

Received : November 18, 2011 Accepted : December 19, 2011

Type of Submission : APJITM Special Corresponding author : Dongwon Lee

1. Introduction

Micro-blogging - a new paradigm of Web-based and mobile communications - has been rapidly growing and gaining explosive popularity worldwide. Compared to traditional online communications, micro-blogging allows more instant and flexible forms of communications. In micro-blogging sites, short length messages (e.g., maximum 140 characters for each message in Twitter) are posted and transferred via a wide variety of communication platforms including the Web on PC, text messaging on mobile phone, instant messenger on smartphone, and other third-party applications. Such a flexible and ubiquitous architecture fundamentally lowers the technological barrier for initiation and further encourages users' continuous use of micro-blogging services.

Also, micro-blogging is now evolving from a daily chatting tool into a significant communication platform where individuals and organizations are able to seek and share real-time news updates upon personal events and social issues. Among micro-blogging sites, Twitter is the most popular one. As of September in 2011, Twitter globally has 100 million active users, and over 600,000 new accounts are added daily. According to Twitter statistics site named Oikolab, Korean Twitter users are more than 5 million in November 2011. As such, Twitter has made remarkable growth in recent years, and its influence on people's daily communications seems to become broader and stronger as time goes by.

Unlike other types of social media, Twitter has several distinctive and attractive features. First, Twitter restricts the length of posted messages to 140 characters. Thus, to communicate others or express their ideas effectively under such limitation, users tend to express their opinion in more direct manners. That makes us understand someone's thoughts more clearly and capture more definite social sentiment (e.g., positive, negative, or neutral) of someone's opinions. Second, due to the distinctive retweet (RT) mechanism, Twitter users are able to share their messages quickly with other users. Moreover, as the use of smartphone becomes more popular, people can access to Twitter more often in any-time and from anywhere. Therefore, users can look up others' opinions or news immediately and easily transfer some useful information to other members.

The power of such rapid message transmission overwhelms traditional news and broadcasting media. Recognizing transformative potentials of Twitter, broadcasting and news corporations in Korea recently begin to utilize Twitter as a significant communication source to supplement shortfalls of their traditional media. For instance, major Korean media corporations such as Chosun, Joongang, Dong-A, SBS, MBC, and KBS also have their own slots on Twitter. Through these Twitter accounts, people can see real-time posted news easily and retweet what they see very quickly. The capabilities of Twitter, in short, enable various personal ideas and opinions rapidly spread throughout digital com-

munity and result in new shape of socio-technological phenomena.

In this paper, we explore RT mechanism by investigating the effect of message characteristics on information diffusion in terms of volume and speed. More specifically, we seek to understand what kinds of messages posted on Twitter have more chance to be retweeted. With these perspectives and ideas in mind, we raise two major research questions; (1) Does RT mechanism take any different pattern depending on diverse characteristics of certain keywords (e.g., the issues of politics, entertainment, celebrities, and products) within posted messages? (2) How much (or how frequently) and how fast are messages for each keyword are transmitted on Twitter?

To identify different RT pattern of each keyword, we classify tweets into more specified categories (i.e., personal opinion, mention, hyperlink, and opinion with hyperlink). And then, we analyze the frequency and velocity of RT for the tweet categories for the certain keyword. We find that more than half of posted messages on Twitter contain personal opinions for the certain keyword, but also find that when being retweeted, the group of messages related to the certain keyword present distinctive diffusion patterns and speed according to message characteristics. Furthermore, the tweets which include objective information with hyperlink are the fastest one. The findings in the study not only broaden our theoretical knowledge of information diffusion mechanism over online social me-

dia but also provide managerial implications regarding how to strategically utilize online social media for marketing communications with customers.

The rest of this paper is organized as follows. In Section 2, we review previous studies on viral opinion and information diffusion and describe Twitter applications and RT mechanism as primary research context. Section 3 presents focal research methodology to examine our research questions. In Section 4, we explain analysis results drawn from the proposed methodology. Finally, we conclude with the discussion of research contributions and future research directions in Section 5.

2. Literature Review

2.1 Viral Opinion and Information Diffusion

The theoretical background related to the current study originates from news diffusion studies that emerged in the early 1960s and were flourished through the 1980s. Against the backdrop of the dominance of mass media, previous research regarded news diffusion as a communication process in which knowledge of a news event spreads through a variety of channels over time across society. Recognizing the audience as a passive receiver, the studies measured diffusion in terms of the awareness or learning of a major news event, such as the assassination of John F. Kennedy or the Challenger disaster (Greenberg 1964). Most of previous news dif-

fusion studies sought to investigate dissemination patterns of information along specific diffusion curves. Scholars in this field primarily focused on examining the way in which a set of factors, including the role of channels, influences information diffusion patterns (Basil and Brown, 1994).

Nowadays, information diffusion studies have increasingly paid more attention to examining when and how people become aware of news through various online communication media, including the Internet (Greenberg et al., 2002) as well as in what patterns they use or adopt new communication channels such as the Internet (Nguyen, 2008). Besides online news diffusion, word-of-mouth (WOM) has been increasingly regarded as an important mechanism through which information can reach large populations, possibly influencing public opinion (Katz and Lazarsfeld, 1995), innovation adoption (Rogers 1995), market share of new products and services, and brand awareness (Keller and Berry, 2003).

In the study of WOM effects, both researchers and practitioners have been interested in investigating whether diffusion of information can be accelerated by a new type of customer group called “influentials” (Weimann, 1994) or “influencers”, who exhibit differentiated personal tendency to proactively communicate opinions or ideas to others by manipulating their expertise and knowledge for products and services. In today’s digital era, the role of these influencers become more emphasized because they

are usually in the center of communications and social networking within the Internet and may significantly affect a disproportionately large number of others (Gladwell, 2000).

Although traditional communication media still remain in dominant place, Web-based social media such as online forums, blogs, and micro-blogs now emerge as alternative forms of rapid dissemination of information (Brownstein et al., 2009). Especially, micro-blogging has been widely used for status updates and live news reports on occasions of emergency where traditional news media cannot be instantly available such as the Southern California wildfires in 2007 (Sutton et al., 2008), the Mumbai terrorist attack in 2008 (Caulfield and Karmali, 2008), the H1N1 Flu outbreak in 2009 (Ostrow, 2009), and the Icelandic volcano eruption in 2010 (Nigam, 2010). Such usages of micro-blogging have received increasing attention from academic researchers.

Hughes and Palen (2009) was among the first studies to investigate the role of micro-blogging in transmitting information under special situations. The study observed Twitter usage patterns surrounding emergency events and compared those with regular use patterns. Hughes and Palen (2009) noted that information propagation was more likely to happen in emergency situations than in regular situations. The study took advantage of the popularity of Twitter and monitored incoming tweets to detect crisis events such as earthquakes and epidemic outbreaks. The research results indicate that the

role of micro-blogging should be expanded from a daily chatting tool into a valuable information sharing platform especially upon emergency events.

Although these kinds of findings are important to understanding the identity and nature of micro-blogging as a salient facilitator of information and news diffusion, prior studies were limited to investigating the relationship between micro-blogging and information diffusion related to the kinds of emergency events. Therefore, the previous research underpinnings provide this study with a chance to explore different types of diffusion patterns of information that is more closely related to people's daily life and their interests.

2.2 Media Characteristics of Twitter

Since its launch in 2006, Twitter has become the largest and most well-known micro-blogging platform. Twitter allows users to send text-based messages (tweets) within maximum 140 characters to a network of followers via a variety of communication channels. By default, tweets are public so that a user can follow and read other users' posts without permission. Twitter is a Web-based communication platform combining instant messaging and SMS that enables subscribers to send short 'status updates' to other people. Because of such features, many researchers have been interested in Twitter as a significant research context. Early studies in this area focused on understanding the

prevalent usage and structural patterns of Twitter. Java et al. (2007) studied the topological and geographical properties of Twitter's social network and summarized different user intentions for using Twitter, such as daily chatting and information sharing. Also focusing on the social networking aspects, Krishnamurthy et al. (2008) characterized distinct classes of Twitter users and their behaviors into broadcasters (e.g., online radio stations and media outlets), acquaintances (e.g., users who exhibit reciprocal relationships), and miscreants (e.g., spammers).

Recent studies have shifted the attention to some novel Twitter applications. For instance, Jansen et al. (2009) studied Twitter as a platform for online word-of-mouth branding. They analyzed more than 10,000 micro-blog posts containing branding information and claimed that micro-blogging could play an important role in designing marketing strategies and campaigns. Ehrlich and Shami (2010) and Zhang et al. (2010) discussed the adoption and use of micro-blogging in the workplace-enterprise micro-blogging. By analyzing users' posting activities and reading behaviors, they found that enterprise micro-blogging could facilitate conversation and mutual assistance. Such user-to-user exchanges and collaborations via Twitter were also identified in a public setting (e.g., Honeycutt and Herring, 2009) in which the authors explored the potential to use Twitter as a collaboration tool. Also, Zhao and Rosson (2009) argued that Micro-blogging has become popular quite quickly, raised its potential for serving as

a new informal communication medium at work, and eventually provided a variety of impacts on collaborative work such as enhancing information sharing, building common ground, and sustaining a feeling of connectedness among colleagues. This exploratory research aims to gain in-depth understandings of how and why people use Twitter and explore micro-blog's potential impacts on informal communication at work.

The rich textual data that are freely available from Twitter also attract interest from the text mining community. O'Connor et al. (2010) applied the sentiment analysis technique to extract public opinions and attitudes from a large body of tweets. They compared the results with opinions derived from standard polling and survey data, which highlighted the promise of using Twitter as a substitute or supplement for traditional polling.

One special characteristic of Twitter is the retweet (RT). It is the key mechanism for online opinion diffusion on Twitter. When Twitter users intend to show a certain tweet to other followers, they can also retweet the contents. Boyd et al. (2010) conducted in-depth study of the various aspects of retweeting. They explicitly interviewed Twitter users and sought to find why the users retweet, and what they retweet the most. Suh et al. (2010) analyzed a large-scale of factors that affect retweeting. The study found that the number of followers and friends significantly affect the number of retweeting, while the number of statuses and favorites do not. Petrovic et al. (2011) more thoroughly inves-

tigated the relationship between influencing factors and retweet patterns in terms of whether, when, how, and how much to retweet. Kwak (2010) found that on the Twitter, there are four main kinds of tweets: singleton, reply, mention, and retweet. The study also found that depending on trendiness of message topics, the portions of these four types of tweet were different.

Although previous literature has examined several aspects of Twitter that originate from its distinctive technological and operational features, the further review of Twitter-related studies indicates that few studies investigate information diffusion patterns on the RT mechanism in terms of volume (e.g., what types of messages are retweeted more frequently?) and speed (e.g., how fast they diffuse?). Therefore, to fill this gap, the current study primarily aims to empirically explore certain diffusion patterns of Twitter information which could be affected by different characteristics of posted messages.

3. Research Methods

3.1 Data Collection

To assess retweeting patterns on Twitter, we collected 91,618 retweeted messages from the Social Metrics for two months from June 1, 2011 to July 31, 2011. Social Metrics is a social trend exploration service Website created by the collaboration between Daumsoft (www.daumsoft.com), a data mining expert firm, and PR Medicom (www.prmedicom.co.kr). <Figure 1> is the snapshot of main screen view of Social Me-



<Figure 1> Screenshots of Social Metrics

tics. This site provides significant Twitter-related statistics including the ranks of popular keywords among posted messages, total amounts of each keyword, other search words closely related to popular keywords, the analysis of social sentiment trends extracted from posted messages, and the daily amounts of messages posted on Twitter for each keyword. Furthermore, the

site also presents information regarding what tweets are relatively more retweeted by other followers with specific numbers. Social Metrics is available for both blog and Twitter, but in this study we used Twitter data and received more detailed data from Daumsoft for the analysis of characteristics of Twitter actual messages.

All sample messages for the current study

are related with one of eight focal keywords including 무상급식, 반값등록금, 나가수, 평창, 김연아, 박태환, 아이폰, and 갤럭시. We believe that the selection of keywords is appropriate for the study because the selected keywords represent popular social aspects of Korea in recent days such as politics, entertainment, celebrity, and technology product. For more details, we selected two keywords in each category; total 8 keywords were included for our analysis. Among many kinds of news category, political and entertainment issues are usually easy to stimulate people's interest, and people usually have more propensities to participate in public debate on these issues. <Table 1> describes summary statistics of sample messages with regard to the selected keywords.

As reported in the <Table 1>, there are

some amounts of tweets not actually related with eight of focal keywords. After excluding these unrelated tweets, total of 81,472 tweets are used for our research. All of these messages are retweeted at least once among Twitter users.

Based on the sample of 81,472 retweeted messages, we calculated the number of Twitter IDs whose messages were retweeted at least once, maximum and minimum frequency of being retweeted by ID, and the maximum frequency of being retweeted by a certain message. <Table 2> summarizes the results described above.

3.2 Research Methodology

In this section, we suggest the methodology to test the retweet (RT) patterns according to keyword characteristics. First of all, to dis-

<Table 1> Summary Statistics of Selected Keywords

	무상급식	나가수	김연아	아이폰	반값등록금	평창	박태환	갤럭시	Total
Number of Tweets	19,615	7,127	5,009	23,740	13,148	11,082	2,330	9,567	91,618
Number of Tweets for Unrelated Topics	4,206	747	598	3,283	26	707	33	546	10,146
Total	15,409	6,380	4,411	20,457	13,122	10,375	2,297	9,021	81,472

<Table 2> Summary Statistics of Sampled Retweeted Messages

Category	무상급식	나가수	김연아	아이폰	반값 등록금	평창	박태환	갤럭시	Total
The number of retweeted ID	1,757	1,955	1,073	5,637	1,599	1,767	629	1,682	5,637
Maximum frequency of retweeted by ID	1,224	300	281	2,659	1,328	794	181	842	2,659
Minimum frequency of retweeted by ID	1	1	1	1	1	1	1	1	1
Maximum frequency of retweeted by certain tweet	406	299	216	438	348	371	108	371	438

cuss what kinds of tweet contents are retweeted more frequently, we divided sampled Twitter messages into specific categories as shown in <Table 3>. So far, very few studies have discussed categorizing of messages posted on Twitter in such categories. Naaman et al. (2010) analyzed what types of messages are commonly posted. Based on their research, we re-categorized each keyword's tweets. Through this category, we classified all of the tweets of each keyword as reported in <Table 3>.

At first, 'personal opinion' refers to the twitter user's subjective thought about a certain keyword. Tweet users can express their own emotion about a topic and criticize some news or others' opinions. Second, 'tweet with link' refers to certain tweets which include objective information with hyperlink. Most of hyperlinks are online news or blogs which include some objective information. On the other hand, 'tweet with link and opinion' includes individual thoughts with hyperlink. Some Twitter users intend to offer hyperlinks for news and information, and add their own opinion about this information. In this case, the tweets are in the category of 'tweet with link and opinion'. And, 'mention'

is Twitter users' reply to other people's tweets. Lastly, the category named 'tweet unrelated topic' means that tweet has certain keyword outwardly, but actually the content is not related to the topic. As mentioned before, we excluded tweets categorized into such 'tweet unrelated topic' from our sample to avoid any semantic and statistical biases that may cause misspecification of information diffusion patterns on Twitter. <Table 4> exemplifies sample messages of each tweet category.

Other than the analysis for whole sample data, we also did the same analysis for the top 100 frequently retweeted messages for each keyword to discover different patterns of retweeting. This is because most of messages on Twitter are usually retweeted just few times. For example, in the sample data, almost 50% of messages are retweeted only at once. While extracting top 100 frequently retweeted messages, we included messages that commonly hold the same rank, resulting in 104 tweets for 무상급식, 111 tweets for 나가수, 116 tweets for 김연아, and 104 tweets for 아이폰.

Next, for analyzing the average speed of each keyword and messages when the certain

<Table 3> Explanation of Tweet Categories

Category	Explanation
Personal opinion	Individual thought about certain keyword
Tweet with link	Objective information with hyperlink
Tweet with link and opinion	Individual thought with hyperlink
Mention	Reply to other followers
Tweet unrelated topic	Tweet including certain keyword, but not related to keyword topic

<Table 4> Examples of Twitter Message Categories

Personal opinion	국민의 80%가 지지하는 사안만 보도하겠다는 매스미디어는 스스로 명을 재촉하는 것이다. 국민의 80%가 지지하는 것은 아마 김연아나 뽀로로 정도밖에 없을 것이다.
Tweet with link	'나가수' 박정현, 올스타전서 애국가 제창(7월 23일 토요일 예정) http://durl.me/cv4j3 TV 중계는 케이블 MBC 스포츠 플러스
Tweet with link and opinion	기어이 하는군요. 무상급식에 반대하는 서울시민 80만여명이 주민투표를 서울시에 청구해 8월 말이면 주민투표가 이뤄질 것 같네요. http://j.mp/jRB80s
Mention	.@doax 어머니가 아이폰 산다고 하시면 전 뜯어말릴 겁니다. 자랑스러운 한국 제품인 삼성의 은하선경이탄이 더 훌륭하거든요.
Tweet unrelated topic	삼성전자 서비스 직원의 불친절에 대해 클레임 걸었더니 삼성전자 서비스 센터 팀장이라는 사람이 전화해서는 대뜸 에어컨 환불은 안된다고 한다. ... 김연아 에어컨 환불 클레임이 엄청 많은가보다 ... 그렇게 잘 줌 만들지 특하면 꺼지는 에어컨을 어케 쓰라고

tweets are retweeted, we consider the tweet which is written during the two months. We conducted the analysis in two ways. The first was measuring the time when an original tweet was first retweeted after being posted, and the second was capturing total time spent for an original tweet to be recursively retweeted within the sample period of the two months between June 1, 2011 and July 31, 2011. Though these methods, we can figure out which keywords and what types of message characteristics are retweeted faster than others.

4. Analysis Results

4.1. Volume of Information Diffusion on Twitter

The goal of the current study, as described in the previous sections, is to explore the dynamics of information diffusion on an online social media environment (i.e., Twitter) from the perspectives of communication volume and speed. To this end, we analyzed more than eig-

hty thousand tweets incorporating eight keywords. For the brief summary of our analysis results, overall research findings in the study indicate that within our sample, RT mechanism shows unique information diffusion patterns depending on message characteristics such as tweet categories.

First, as reported in <Table 5>, we found that for all eight keywords, personal opinion type messages were most frequently retweeted commonly. The unique feature of Twitter explains the result. Due to the limitation of the length of posted messages, Twitter users express their own thoughts using more straightforward expressions. Such way of writing messages usually makes opinion-type tweets more impressive and persuasive to readers, more easily prompting them to retweet the messages.

Next to personal opinion, the category of 'tweet with link and opinion' takes second largest portion of retweeted messages except for the keywords of '나가수', and '김연아.' For these two keywords, messages in 'tweet with link' category turned out to be more frequently re-

tweeted than those in ‘tweet with link and opinion’ category. The reason for such results would be that the two keywords represent recent popular entertainment and sports celebrity for which most of related messages contain or are linked with video clips, and people more tend to share the video clips instead of adding their own opinions. Lastly, the lowest portion of category is ‘mention’ through all of keywords.

<Table 6> presents retweeting portion across message categories among Top 100 tweets. In this part, we analyzed top 100 tweets which were retweeted most frequently with one keyword

of each social issue (e.g., ‘무상급식’, ‘나가수’, ‘김연아’, and ‘아이폰’). Similar to the analysis for the whole sample, the category of ‘personal opinion’ takes the greatest portion and ‘reply’ has the least portion as well. Other message categories show similar portions to the analysis of total sample. The results shown in both <Table 5> and <Table 6> imply that unlike traditional communication media that deliver news and information unilaterally, micro-blog communications like Twitter provide more plausible environment where people’s thoughts on social issues can be easily transmitted to the public.

<Table 5> Content Classification of Selected Keywords(All Tweets)

Category(%)	무상급식	반값등록금	나가수	평창	김연아	박태환	아이폰	갤럭시
Personal opinion	11,460 74.4%	9,589 73.1%	4,843 75.9%	6,291 60.6%	3,187 72.3%	1,351 58.8%	11,442 55.9%	4,260 47.2%
Tweet with link	1,380 9.0%	703 5.4%	794 12.4%	270 2.6%	800 18.1%	269 11.7%	2,804 13.7%	1,160 12.9%
Tweet with link and opinion	2,394 15.5%	2,391 18.2%	633 9.9%	3,724 35.9%	387 8.8%	664 28.9%	5,703 27.9%	3,204 35.5%
Reply (mention)	175 1.1%	439 3.3%	110 1.7%	90 0.9%	37 0.8%	13 0.6%	508 2.5%	397 4.4%
Total	15,409	13,122	6,380	10,375	4,411	2,297	20,457	9,021

<Table 6> Content Classification of Each Keyword(Top 100 Tweets)

Category(%)	무상급식	나가수	김연아	아이폰
Personal opinion	77 (66.3%)	73 (57.7%)	71 (49.1%)	55 (41.3%)
Tweet with Link	7 (6.7%)	8 (7.2%)	14 (12.1%)	6 (5.8%)
Tweet with Link and opinion	16 (15.4%)	7 (6.3%)	8 (6.9%)	37 (35.6%)
Reply (mention)	0 (0.0%)	1 (0.9%)	0 (0.0%)	2 (1.9%)
Tweet unrelated topic	4 (3.8%)	22 (19.8%)	23 (19.8%)	4 (3.8%)
Total	104	111	116	104

4.2 Speed of Information Diffusion on Twitter

To understand the other aspect of retweet mechanism, we need to configure how fast Twitter messages are disseminated across users over time. We analyzed the velocity of Twitter diffusion in two ways. First, we calculated the time gap between the moment of an original tweet being posted and then being retweeted first (which is called ‘First RT’). Second, we captured the total time spent for an original tweet to be continuously transmitted (i.e., re-tweeted) within two months sample period (which is referred to as ‘Total RT’). We also examined both ‘First RT’ and ‘Total RT’ in terms of four categories of Twitter messages.

According to the results in <Table 7>, we understand that First RT is quite different to each keyword characteristic. As a result of first analysis, the average First RT of ‘무상급식’ and ‘반값등록금’ which refer to political issues are the fastest. The result implies that people should be more interested in political issues in

Twitter space. We can find the similar result in Yu et al. (2008)’s research. They found that the average sentiment level of congressional debate was higher than that of neutral news article. Sentiment level is related to the amount of tweets and buzz of its topic.

The celebrity category which includes ‘김연아’ and ‘박태환’ shows the second fastest time for First RT. The First RT for messages with the keywords of ‘나가수’ and ‘평창’ is relatively slow. The interesting thing is that messages with ‘아이폰’ and ‘갤럭시’ keywords show different results although they are in the same ‘product’ category. The First RT for messages with ‘갤럭시’ is slower than that for messages with ‘아이폰’. However, messages with the keyword of ‘아이폰’ are retweeted more slowly over time. This is because ‘아이폰’ just represents a single product and thus discourages users from transmitting related messages. However, ‘갤럭시’ includes many kinds of family products such as ‘갤럭시 탭’ and ‘갤럭시 s2’, providing users with more something to talk

<Table 7> RT Speed of Each Keyword(Unit : Minute(s))

	무상급식		나가수		김연아		아이폰	
	First RT	Total RT	First RT	Total RT	First RT	Total RT	First RT	Total RT
Min.	0	0	0	0	0	0	0	0
Max.	971	37,990	23,344	50,264	3,030	60,480	1,467	54,250
Avg.	14	1,404	218	1,218	40	1,763	23	1,835
	반값등록금		평창		박태환		갤럭시	
	First RT	Total RT	First RT	Total RT	First RT	Total RT	First RT	Total RT
Min.	0	0	0	0	0	0	0	0
Max.	1,399	69,499	12,219	17,635	1,217	15,536	4,982	48,715
Avg.	29	1,492	114	787	32	437	83	944

about and motivation to share information.

Repeatedly, Total RT refers to the total time spent from the time of first being re-tweeted to that of lastly being re-tweeted. As shown in the <Table 7>, messages with ‘김연아’ and ‘아이폰’ are slowly re-tweeted for a longer period of time (i.e., average Total RT for

these keywords is larger). The results indicate that many people talk about these topics all the time. The average Total RT for ‘무상급식’, and ‘반값등록금’ are the fastest. The reason for the result is that this political issue has been perceived as critical events and then many people would like to raise and share their own opinions.

<Table 8> RT Speed of Each Keyword by Message Characteristics(Unit : Minute(s))

		무상급식		나가수		김연아		아이폰	
		First RT	Total RT	First RT	Total RT	First RT	Total RT	First RT	Total RT
Personal opinion	Min.	0	0	0	1	0	0	0	0
	Max.	233	37,990	1,171	50,264	277	60,480	751	50,275
	Avg.	13	1,740	42	1,865	11	2,284	10	1,514
Link	Min.	0	23	0	0	0	1	0	3
	Max.	41	3,677	887	27,698	241	15,117	529	54,250
	Avg.	7	75	58	1,096	18	1,495	17	2,418
Reply	Min.	0	0	0	373	0	109	0	1
	Max.	32	1,605	23,344	2,133	402	7,795	925	1,968
	Avg.	9	228	1,246	81	122	1,472	101	241
Link +opinion	Min.	0	30	0	1	0	8	0	1
	Max.	971	6,411	4,975	4,265	3,030	7,336	1,467	22,886
	Avg.	28	1,188	215	334	149	812	30	2,076
		반값등록금		평창		박태환		갤럭시	
		First RT	Total RT	First RT	Total RT	First RT	Total RT	First RT	Total RT
personal opinion	Min.	0	0	0	0	0	0	0	0
	Max.	874	69,499	12,219	12,309	1,217	2,556	944	48,715
	Avg.	20	2,473	161	819	36	311	40	994
Link	Min.	0	1	0	0	0	0	0	4
	Max.	476	3,077	85	17,635	39	1,691	150	7,387
	Avg.	19	355	11	1,208	6	319	18	374
Reply	Min.	1	2	0	51	0	0	0	0
	Max.	902	662	2,266	258	330	221	4,982	8,326
	Avg.	66	83	127	25	40	19	371	494
Link +opinion	Min.	0	1	0	0	0	2	0	0
	Max.	1,399	19,953	841	16,996	512	15,536	4,540	41,525
	Avg.	31	967	43	856	33	829	118	1,295

For more detailed analysis, as shown in <Table 8>, we recalculated both First RT and Total RT for focal keywords in terms of messages categories. It was for the purpose of increasing our understanding of what types of Twitter messages are diffused faster than others and how each keyword leads to different retweeting patterns. Unlike the analysis of retweeted volume, ‘personal opinion’ is not the category which is retweeted fast. Instead, ‘tweets with link’ are retweeted most rapidly. Next to ‘tweet with link’ category, tweets including both objective information link and subjective personal opinion are the second fastest.

For the summary, as shown in <Table 9>, the gap between First RT and Total RT is relatively small for messages of ‘tweet with link’ category. Consistent with the result in <Table 8>, ‘tweet with link’ category holds the fastest First RT, which implies that Internet technologies such as hyperlink may help facilitating diffusion of online messages on social media environments. Also, even from aggregated information of 8 keywords total, average Total RT for messages in ‘personal opinion’ category is the slowest. The results infer that many users

tend to more frequently retweet ‘personal opinion’ type tweets for a long time.

Additionally, we examined the role of user types in dispatching information on Twitter for top 100 tweets. The analysis aims to identifying what types of users significantly contribute to retweeting specific messages containing any of focal keywords. We classified Twitter users into six categories; news media, celebrity, company, general public, government, and other organizations and calculated number messages retweeted by each category of user group among top 100 tweets. The results are summarized in <Table 10>. For the messages with ‘반값등록금’, ‘김연아’, and ‘박태환’, almost half of Top 100 tweets are retweeted by general public. Celebrity and news media also contribute to contribute to retweeting messages at certain levels.

However, our results indicate that company users do not sufficiently affect re-transmitting Twitter messages. This is because many of keywords for this study are rather about political and entertainment issues than corporate advertisements and product information. The results imply that online social media like Twitter should be favorable environment for gathering

<Table 9> RT Speed of Each Category by Message Characteristics(Unit : Minute(s))

		8 Keywords total	
		First RT	Total RT
personal opinion	Average	171	10,471
Link	Average	22	918
Reply	Average	323	442
Link +opinion	Average	95	1,053

<Table 10> ID Classification of Each Keyword(Top 100 tweets)

Subjects of retweeting	무상급식	반값등록금	나가수	평창	김연아	박태환	아이폰	갤럭시
Media and Press	21	14	19	19	19	19	12	15
Celebrity	36	23	39	19	13	12	28	17
Company	1	1	10	16	16	15	15	24
Public	41	50	36	34	51	49	39	38
No information	4	10	4	8	6	9	5	5
Government	0	0	0	5	3	4	1	0
Organization	1	2	1	1	0	1	0	1
Total	104	100	111	102	116	109	104	100

information and ideas generated by general public who share common interest in personal or social affairs.

5. Discussion and Conclusion

In this paper, we examined the effects of message characteristics on information diffusion mechanism under social media environments. Based on the research context of Twitter, we empirically analyzed the volume and speed of retweeting for posted messages from June 1, 2011 to June 31, 2011.

Focusing on the message and keyword features, we found that each keyword had both common and differentiated effects on message transferring across members of Twitter. In volume wise, through the all focal keywords, ‘personal opinion’ type messages were retweeted overwhelmingly. The findings indicate that in micro-blogging site such as Twitter, the public has the power to deliver and diffuse specific information with their own thought to other people easily. For these results, we propose that the

underlying reason is one of the unique feature of Twitter; limited length of tweets which usually makes messages more impressive and straightforward. Leveraged by such feature, personal opinions tend to be easily accepted by other Twitter users

In the aspect of information diffusion speed, we sought to understand what types of tweets are more swiftly transmitted and how each keyword comes up with different retweeting pattern respectively. As a result, tweets with objective hyperlink turned out to be disseminated most rapidly. Also, we found that the average total RT speed of political issues such as ‘무상급식’, and ‘반값등록금’ were the fastest. For such phenomenon, it can be explained that political issue is usually prone to be considered a big social event, and many people want to present their own opinions on the open discussion space easily accessible to the public.

Aggregating the research findings in the current study, we argue that on social media environments, features of messages including keywords, content, and types of authors play piv-

otal role in configuring information diffusion patterns. As an exploratory study, the research findings in the current study provide both academic and practical contributions. From the academic perspective, the study broadens our theoretical knowledge of information diffusion mechanism over online social media. Also, the study provides novel research perspectives such as concepts of volume and speed of information diffusion, which have been rarely investigated but are useful for understanding the dynamics and nature of online social media. For practitioners, the research findings also offer managerial implications regarding how to strategically utilize online social media for marketing communications with customers.

However, the study has some limitations that should be addressed in the future study. First, limited number of message keywords applied to the study makes it hard to generalize the research findings. Thus, in order to enhance the applicability of the study to broader range of real world environment, additional efforts need to be devoted to analyze information diffusion patterns of more diversified message categories that represent various social aspects. Second, the study applied relatively simple methods for empirical analysis, and thus the research findings still remain at the summary statistics level. To derive more enriched theoretical and managerial implications, more sophisticated methodologies such as data mining, experimental analysis, and econometric analysis should be implemented for the future study.

References

- Basil, M. and W. Brown, "Interpersonal Communication in News Diffusion : A Study of 'Magic' Johnson's Announcement", *Journalism Quarterly*, Vol.71, No.2(1994), 305 ~ 320.
- Boyd, D., S. Golder, and G. Lotan, "Tweet, Tweet, Retweet : Conversational Aspects of Retweeting on Twitter", in *Proceedings of the 43th Hawaii International Conference on System Sciences(HICSS)*, 2010.
- Brownstein, J. S., C. C. Freifeld, and L. C. Madoff, "Influenza A (H1N1) Virus, 2009-Online Monitoring", *New England Journal of Medicine*, Vol.360(2009), 2156.
- Caulfield, B. and N. Karmali, "Mumbai : Twitter's Moment", *Forbes.com*, 2008.
- Ehrlich, K. and N. S. Shami, "Microblogging Inside and Outside the Workplace", in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media(ICWSM)*, 2010.
- Gladwell, M., *The Tipping Point : How Little Things Can Make a Big Difference*, Little Brown, New York, NY, 2000.
- Greenberg, B. S., "Diffusion of News of the Kennedy Assassination", *Public Opinion Quarterly*, Vol.28, No.2(1964), 225 ~ 232.
- Greenberg, B. S., L. Hofschire, and L. Lachlan, "Diffusion, Media Use and Interpersonal Communication Behaviors", in Greenberg, B. S. (Ed.), *Communication and Terrorism*, Hampton Press, Cresskill, NJ, (2002), 3 ~ 16.
- Honeycutt, C. and S. C. Herring, "Beyond Microblogging : Conversation and Collaboration via Twitter", in *Proceedings of the 42nd Hawaii International Conference on System Sciences(HICSS)*, 2009.

- Hughes, A. L. and L. Palen “Twitter Adoption and Use in Mass Convergence and Emergency Events”, *International Journal of Emergency Management*, Vol.6, No.3/4(2009), 248 ~ 260.
- Jansen, B. J., M. Zhang, K. Sobel, and A. Chowdury, “Micro-Blogging as Online Word of Mouth Branding”, in *Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems (CHI)*, (2009), 3859 ~ 3864.
- Java, A., X. Song, T. Finin, and B. Tseng, “Why We Twitter : Understanding Microblogging Usage and Communities”, in *Proceedings of the 9th WEBKDD and 1st SNA-KDD Workshop on Web mining and Social Analysis*, 2007.
- Katz, E. and P. F. Lazarsfeld, *Personal Influence : The Part Played by People in the Flow of Mass Communications*, Free Press, Glencoe, IL, 1955.
- Keller, E. and J. Berry, *The Influentials : One American in Ten Tells the Other Nine How to Vote, Where to Eat, and What to Buy*, Free Press, New York, NY, 2003.
- Krishnamurthy, B., P. Gill, and M. Arlitt, “A Few Chirps about Twitter”, in *Proceedings of the 1st Workshop on Online Social Networks(WOSN)*, (2008), 19 ~ 24.
- Kwak, H., C. Lee, H. Park, and S. Moon, “What is Twitter, a Social Network or a News Media?”, in *Proceedings of the WWW 2010 Conference*, (2010), 26 ~ 30.
- Naaman, M., J. Boase, and C.-H. Lai, “Is it Really About Me? Message Content in Social Awareness Streams”, in *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*, (2010), 6 ~ 10.
- Nguyen, A., “The Contribution of Online News Attributes to Its Diffusion : An Empirical Exploration Based on a Proposed Theoretical Model for the Micro-Process of Online News Adoption/Use”, *First Monday*, Vol. 13, No.4(2008).
- Nigam, S., “How Social Media Helped Travelers During the Iceland Volcano Eruption”, *Mashable Business*, 2010.
- O’ Connor, B., R. Balasubramanian, B. Routledge, and N. Smith, “From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series”, in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media(ICWSM)*, 2010.
- Ostrow, A., “Swine Flu Hysteria : 10,000 Tweets Per Hour”, *Mashable Social Media*, 2009.
- Petrovic, S., M. Osborne, and V. Lavrenko, “RT to Win! Predicting Message Propagation in Twitter”, in *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*, 2011.
- Rogers, E. M., *Diffusion of Innovations* (4th Ed.), Free Press, New York, NY, 1995.
- Suh, B., L. Hong, P. Pirolli, and E.H. Chi, “Want to Be Retweeted? Large Scale Analytics on Factors Impacting Retweet in Twitter Network”, in *Proceedings of the IEEE 2nd International Conference on Social Computing*, (2010), 177 ~ 184.
- Sutton, J., L. Palen, and I. Shloviski, “Back-Channels on the Front Lines : Emerging Use of Social Media in the 2007 Southern California Wildfires”, in *Proceedings of the 2008 ISCRAM Conference*, Washington, DC, 2008.
- Weimann, G., *The Influentials : People Who Influence People*, State University of New York Press, Albany, NY, 1994.

- Yu, B., S. Kaufmann, and D. Diermeier, “Exploring the Characteristics of Opinion Expressions for Political Opinion Classification”, in *Proceedings of the 9th Annual International Conference on Digital Government Research*, 2008.
- Zhang, J., Y. Qu, J. Cody, and Y. Wu, “A Case Study of Micro-Blogging in the Enterprise : Use, Value, and Related Issues”, in *Proceedings of the 28th International Conference on Human Factors in Computing Systems(CHI)*, (2010), 123~132.
- Zhao, D. and M. B. Rosson, “How and Why People Twitter : The Role that Microblogging Plays in Informal Communication at Work”, in *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, 2009.

Abstract

트위터 메시지 특성에 따른 온라인 구전효과에 대한 분석

남영우* · 손인수* · 이동원*

본 연구에서는 최근 온라인 소셜미디어 서비스로 각광받고 있는 트위터 상에 게재되는 메시지의 특성이 메시지의 재전달에 미치는 영향을 전달 빈도와 전달 속도 관점에서 분석함으로써 소셜네트워크 환경 하에서의 정보확산 체계를 탐구하고자 하였다. 이를 위해 최근 소셜미디어 상에서 많이 언급되고 있는 정치, 연예, 스포츠, 유명인, 최신 디지털 제품과 같은 사회적 관심사를 나타내는 여덟 가지 주요 키워드(예 : 무상급식, 반값등록금, 나가수, 평창, 김연아, 박태환, 아이폰, 갤럭시)을 선정하였다. 또한 트위터 메시지의 특성을 정의하기 위한 기준은 기존 문헌연구를 통해 개인의견, 하이퍼링크 포함 메시지, 하이퍼링크 및 개인의견 포함 메시지, 단순 재인용 등의 네가지 유형을 적용하였으며 이들 주요 키워드와 메시지 유형별로 트위터 메시지의 재전달 빈도와 재전달 속도를 측정하였다. 2011년 6월부터 2011년 7월까지 수집된 81,472개의 1회 이상 재전달된 트위터 메시지에 대한 분석을 통해, 재전달 트위터 메시지의 50퍼센트 이상이 특정 키워드에 대한 개인의견을 표현한 메시지임을 알 수 있었다. 아울러 표본 트위터 메시지 중 하이퍼 링크를 포함한 메시지가 가장 빠른 메시지 재전달 시간을 보이는 것을 알 수 있었다. 이러한 분석결과를 바탕으로 메시지의 고유 특성이 온라인 소셜 미디어 상의 정보확산 패턴에 중요한 영향요소로 작용하고 있음을 확인 할 수 있었다. 또한 트위터와 같은 온라인 소셜 미디어는 더 이상 단순한 의사소통 도구가 아닌 다양한 사회현상에 대한 사용자의 개인적 의견을 피력하고 상호간의 심도있는 논의가 가능한 현대사회의 중요한 커뮤니케이션 플랫폼으로 발전하고 있으며 정보기술이 이러한 의사소통 혁신에 중요한 요소로 작용하고 있다는 것을 이해 할 수 있었다. 본 연구의 결과는 온라인 소셜 미디어상의 정보확산에 대한 기존의 학문적 지식을 증대시키고 실무적으로 기업이 고객과 소통하는 데 있어 온라인 소셜 미디어를 어떻게 전략적으로 활용할 수 있는가에 대한 실질적인 방향을 제시할 것이라 기대된다.

Keywords : 정보확산, 메시지 특성, 소셜미디어, 트위터

* 고려대학교 경영대학

저 자 소개



남영우

고려대학교 경영대학 대학원에서 MIS 석사과정에 재학 중이다. 관심분야는 소셜 미디어의 확산 및 사용 패턴이다.



손인수

현재 고려대학교 경영대학 대학원 박사과정 재학 중이다. 고려대학교 경제학사, University of Wisconsin-Madison 경영학 석사(MBA), Carnegie-Mellon University 경영정보학 석사 (MISM)를 취득하였으며 LG-CNS 컨설팅 부문에서 근무하였다. 주요 관심 연구분야는 E-business, Service-oriented computing, Technology-based innovation, Business value of IT, Consumer behaviors under social network environment이다.



이동원

현재 고려대학교 경영대학 교수로 재직 중이다. 서울대학교에서 경영학사 및 경영학 석사, University of Arizona에서 MIS 석사, University of Minnesota에서 경영학 박사를 취득하였다. 주요 관심 연구분야는 전자상거래 가격 전략, 온라인 소비자 행동 분석, 소셜미디어, 클라우드 컴퓨팅 등이다. Review of Economics and Statistics, Journal of AIS, Journal of MIS, International Journal of Electronic Commerce, Communications of the AIS, Information Systems Frontier, Electronic Markets, Journal of Global Information Technology Management, ACM Crossroads 등을 비롯한 다수의 국내외 학술지에 논문을 발표하였다.