

Content-Based Analytics of Diffusion on Social Big Data: A Case Study on Korean Telecommunication Companies

Namhee Lee and Jason J. Jung^(✉)

Department of Computer Engineering, Chung-Ang University,
Seoul 156-756, South Korea
jjjung@gmail.com

Abstract. Social networking services have been playing an important role of communicating with customers. Particularly, firms seek to deploy Twitter for the benefit of their business because it has rapidly become an information vehicle for consumers who are disseminating information on products and services. Thus, this study examines how information shared by firms is diffused and what the important factors in understanding information dissemination are. Specially, this study classifies the types of tweets posted by a firm (@olleh_mobile) and then to investigate the effect of these types of tweets on diffusion. By using content analysis, this study defined two categories ('Information providing' and 'Advertisement' type) and eight subordinate concepts (News, Usage, Preview, Notice, Sale, Benefit, Event, Service public relations). These results indicate that the differences are significant for all three types of information content. It shows that firms can spread information more quickly by providing the 'Information and advertisement' type rather than the 'Advertisement' type.

Keywords: Information contents · Information diffusion · Information types · Twitter · Content-based analytics

1 Introduction

Social media and communications technology have become important drivers for new types of communication and have made users better able to share information (Smith 2010). New communication channels and mechanisms like Facebook, Twitter and Wikis, allow the creation and exchange of content that has been created by users and provide collaborative structures where user interaction is proactively encouraged (Kaplan and Haenlein 2010). New media and communication platforms make it easier to spread information very quickly and make one person able to communicate with hundreds or thousands of others.

One well known social media channel, Twitter, is a short message service. It has attracted advertising and marketing interest from firms to improve brand trust and loyalty for customers (Sledgianowski and Kulviwat 2009). Twitter allows firms to engage in timely and direct end-user contact at relatively low cost and higher levels of efficiency than can be achieved with more traditional communication tools (Kaplan and

Haenlein 2010). Therefore firms cannot ignore the importance of Twitter because it has rapidly become an information vehicle for consumers who are disseminating information on products and services (Fischer and Reuber 2011).

As a growing number of firms seek to deploy Twitter for the benefit of their business, the current study extends investigations that consider how companies use Twitter to facilitate dialogic communication with their consumers (Greer and Ferguson 2011; Jansen et al. 2009; Li and Rao 2010). Even though there are various discussions about how information can be disseminated on Twitter, most research topics on Twitter are related to personal, social or public news (Cha et al. 2010; Li and Rao 2010). This scope of this study is the type of determinants influencing information diffusion for firms which use Twitter. Thus, in spite of the growth of Twitter, the business viability of Twitter remains in question. Also, managers within firms are still uncertain as to how Twitter can be used in marketing and which types of message exert the most influence or get reposted by users (Kim et al. 2012).

Therefore, this study examines how information shared by firms is diffused and what the important factors in understanding information dissemination are. More specifically, this study poses the research questions: (1) What types of information are provided by a firm using Twitter? (2) How does the diffusion of information that is posted by a firm differ by different types of information? Through these questions, this study will suggest content that is appropriate for diffusion and the relationship between patterns of tweets and diffusion on Twitter. It is expected that there can be a significant impact derived from analysis of information diffusion to guide firms Twitter use.

The remainder of this study is structured as follows. In the next section, we present the theoretical background and develop the research framework. This is followed by a description of the methods employed, including the data, variable operationalization, and analysis techniques. Finally, this study presents results, and ends with a discussion of limitations and theoretical and practical implications of the findings.

2 Theoretical Background

2.1 Understanding Twitter

Twitter is an online social networking and micro-blogging service. It is a social networking service because users have a profile page and connect to other users by following them (Thelwall et al. 2011). Also, Twitter offers a micro-blogging service by allowing its users to send messages- called “tweets” – to their followers while visiting other users’ accounts (Savage 2011). Tweets are text-based posts of up to 140 characters in length. On Twitter, users can post original tweets under their Twitter accounts and can “Retweet”, which means posting another user’s tweet. When a person chooses to follow someone, they receive their tweets (Fischer and Reuber 2011). The purpose of retweeting is to diffuse information to followers, and this diffusion seems to be extremely rapid (Thelwall et al. 2011). Therefore, retweeting is the key mechanism for information diffusion in Twitter.

2.2 Previous Research of Twitter Usage by Business

Many firms use Twitter as a marketing tool. Twitter is useful to disseminate information or messages related to their products or service to customers. A number of firms use Twitter to disseminate information to stakeholders (Jansen et al. 2009). Also, Twitter gives business an opportunity to track what consumers are saying about their products and allow consumers to post instant opinions about a certain brand even though not have any previous relationship (Savage 2011). According to a recent report, 103 Chief Marketing Officers responded to the question asking which platform would figure into their marketing plans the most in the coming months. 40.8 % responded Twitter, followed by 26.2 % saying Facebook, 16.5 % saying LinkedIn, and 8.7 % responding “Other”. For example, Dell Outlet used Twitter to reach consumers and found out that consumers were interested in communicating via Twitter.

Research of business usage of Twitter is at a very early stage of development (Barnes and Böhringer 2011). Zhang and Watts (2008), Jansen et al. (2009), and Berinato (2010) showed how to use Twitter for promotion and branding in firms. In these studies, Twitter is described as a tool to create communication with consumers. While previous work on Twitter has been extensive, it has generally focused only on single perspectives or factors for using Twitter.

2.3 The Diffusion of Innovations Theory

Diffusion theory has been studied in a variety of contexts and from many perspectives. The Rogers model provides a reasonably comprehensive view of innovation diffusion (Brancheau and Wetherbe 1990). Since the publication of Rogers’s widely referenced work, the diffusion model’s focus on the individual adoption process and emphasis on communication behavior have been extended to technology and information adoption (Bajwa et al. 2008; Chatman 1986).

Rogers (1983) defines diffusion to be “the process by which an innovation is communicated through certain channels over time among the members of a social systems”. So diffusion is a special type of communication in which the messages are about new ideas (Rogers 2003). Innovation, as an idea or object, is a key factor in diffusion theory. The characteristics of innovation are one important explanation of diffusion. Rogers and Shoemaker (1971) identify the most important innovation characteristics that influence the adoption of an innovation as trialability, relative advantage, compatibility, observability, and complexity. Previous studies have considered innovation or information characteristics such as usefulness, accuracy, and source or credibility (Cheung et al. 2008). These characteristics influence the diffusion of innovations or information about new ideas.

Thus, the diffusion approach helps to understand how individuals behave as they consider the adoption of an innovation. Moreover, it is a useful theory to know the diffusion process in online communities. The characteristics of information or messages in social networking may play a special role in diffusion (Zhao and Rosson 2009). However, research on diffusion has concentrated primarily on innovation and its characteristics (Chatman 1986). Research perspectives are limited to understanding the diffusion patterns of information.

3 Research Model and Hypotheses

Based on classical diffusion of innovation theory, this study developed a research model as shown in Fig. 1. Consistent with this purpose of this study, information content in tweets provided by firms are classified into three types. The model examines the effect of patterns of information content on three attributes of information diffusion (scale, speed and duration).

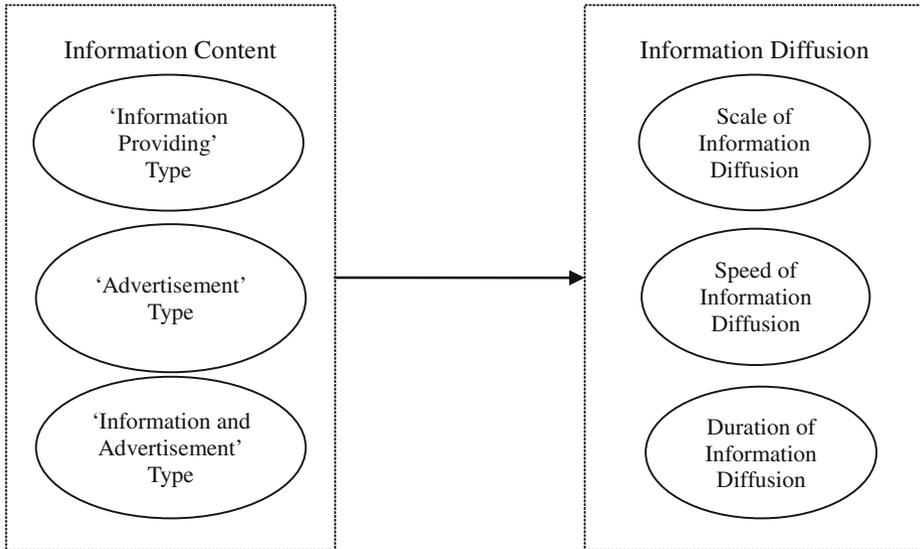


Fig. 1. The conceptual research model

Information characteristics influence a user's decision to spread information. Individual perceptions of the attributes of information can affect its rate of adoption (Rogers 1983; Rogers 2003), Chatman (1986) suggests that the phenomenon of diffusion can shift from information characteristics or contents. For example, information about a job can be circulated quickly because this type of information is time-sensitive. Most studies focus on characteristics of innovation and information such as usefulness, message relevance, accuracy, comprehensiveness (Cheung et al. 2008). It measures how users perceive these characteristics as related to its value, and in relationship to their past experience (Rogers 2003).

Even on Twitter, the types of information tweets suggest is important to diffusion. Tweets can be categorized based on many types of content. Java et al. (2007) identify tweets' contents with reference to information sharing, information seeking, and relationships. Naaman et al. (2010) suggest nine content categories including information sharing, opinion, complaint, self-promotion etc. The characteristics of messages shared on Twitter may play a special role in sharing or delivering information (Zhao and Rosson 2009). Consequently, information diffusion will depend on the type of content.

Hypothesis 1

The content of information will have a significant effect on information diffusion.

Hypothesis 1a

The content of information will have a significant effect on the scale of information diffusion.

Hypothesis 1b

The content of information will have a significant effect on the speed of information diffusion.

Hypothesis 1c

The content of information will have a significant effect on the duration of information diffusion.

4 Research Methodology

This study was conducted to test the proposed model for information diffusion on Twitter. The research model was applied to the content and type of information in tweets created by a firm. A content analysis of tweets was performed using coding messages. Then a cluster and non-parametric ANOVA analysis was conducted to determine the types of information content that influence diffusion.

4.1 Data Collection

As the study's objectives are first to classify the types of tweets posted by a firm and then to investigate the effect of these types of tweets on diffusion, an in-depth perspective such as that offered by the Twitter experiences of a particular firm was deemed appropriate. In order to collect a realistic Twitter dataset, this study used a commercial system (called TweetScope). This system can manage tweet streams from multi-user accounts simultaneously in real-time. Using TweetScope, 1006 tweets posted by the @olleh_mobile account, from March 16th, 2012 to October 11th, 2012 were downloaded. The Olleh_mobile business has been actively interested in Twitter to promote the brand. It has 116,376 followers, 496 followings and 105,414 tweets. Olleh_mobile has posted a total of 104,824 tweets and 25 tweets daily on average. The data includes the content of tweets, tweet time, and the number of retweets (RTs). Also, this data contains information on users who repost tweets, user ID, times of users' RT, the number of followers, followings, and total tweets. Excluded were mention tweets because these are not delivered automatically to followers. This study analyzed 1006 tweets.

4.2 Measure

4.2.1 Dependent Variables: Information Diffusion

In Rogers (1983) model, diffusion was defined as how many people use or adopt new ideas, or technology and measured the cumulative number of adopters. Previous studies measured diffusion by means of dependent variables such as binary adoption/non-adoption, time of adoption, and frequency of use (Fichman 1992).

Table 1. Dependent variables definitions

Variables	Definition	Measure	Reference
The scale of RTs	How many times a tweet is reposted by users	The number of RTs	Cheung et al. (2008)
The speed of RTs	How fast a tweet is reposted by users	Total RTs (n)/Starting time of RT – Ending time of RT (hour)	Yang and Counts (2010)
The duration of RTs	How long a tweet is reposted by users	The gap between Starting time of RT and Ending time of RT (hour)	

Similarly, on Twitter, diffusion means how much or far tweet propagates throughout the community of users or followers. Diffusion is generally measured as the number of retweets – these are posted tweets sent to followers on Twitter. Yang and Counts (2010) developed three dimensions of diffusion: speed, scale and range. In this study, information diffusion is measured by three dimensions, as shown in Table 1.

4.2.2 Independent Variables: Information Contents

4.2.2.1 Coding Framework and Procedure

This study used content analysis by first conceptualizing the content of tweets to characterize the rationale for tweets posted by a firm on Twitter. Recently, as research about social media has increased, much of it focuses on the purpose of information or communication and also on new communication behaviors (Herring et al. 2004; Papacharissi 2004). This research uses content analysis. This analysis was conducted to identify and quantify structural factors or properties of websites, blogs and social network sites (Ha and E.L. James 1998; Lin and Pena 2011).

Content analysis is considered to be a qualitative research method. It is effective in examining both theoretical definitions and empirical measurements. This analysis provides researchers with opportunities to unobtrusively study the values, sentiments, intentions and ideologies of sources generally inaccessible to researchers (Morris 1994). The goal of content analysis is to create objective criteria for transforming written text contained in highly reliable data (Simmons et al. 2011). It classifies content analysis methods into three types: (1) human-scored schema, (2) individual word count systems, and (3) computerized systems using artificial intelligence. Content analysis may provide an effective tool for gaining access to desired study information.

In this study, to analyze the content of tweets, two coders independently looked through downloaded tweets and classified the purpose of tweets based on previous studies and theories. Then the study compared and contrasted each classification deduced by coders.

In this step, while two coders agree on representations for three types of content and different types of subordinate categories, while also agreeing on the elimination of inappropriate concepts. Finally, tweets provided by a firm were classified into two categories (information providing and advertisement) and eight subordinate content

Table 2. The level of agreement in information content

Constructs	Variables		The level of agreement	N
Information contents	Information providing	News	86 %	58
		Usage	93 %	336
		Review	97 %	225
		Notice	84 %	145
	Advertisement	Sale	85 %	29
		Benefit	89 %	199
		Event	92 %	159
		Service public relations	92 %	46

groupings (News, Usage, Review, Notice, Sale, Benefit, Event, Service PR) as shown below in Table 2.

Two coders received training and analyzed the tweets according to the predetermined categories. Each coder had the chance to practice the procedure with a few examples before the actual coding started. A tweet could be classified into more than one category. For example, a tweet that provides information could also express advertisement. Overall intercoder reliability, as measured by their performance on the total set of tweets, was assessed by using Cohen's kappa. Cohen's kappa is a more rigorous means of assessing reliability than using other statistics such as an exact percent agreement because it accounts for chance when measuring the level of agreement between two coders (Boettger and Palmer 2010). For this study the kappa test identified an overall agreement of above 80 %, indicating an acceptable level of consistency between coders.

4.3 Types of Information in Firm's Twitter Messages

A few researches looked at objectives of social media by using content analysis. Fullwood et al. (2009) categorize the purpose of blogs into categories diary, advertising, providing information, sharing media, emotional outlet, and reporting. Naaman et al. (2010) developed a content-based categorization of messages posted by Twitter users. They suggested nine message categories: information sharing, self promotion, opinion, statements, and questions to followers, presence maintenance, and anecdotes. Furthermore, Java et al. (2007) identified four user objectives on Twitter, daily chatter, conversations, sharing information and reporting news.

Based on the knowledge gained from these studies, this study defined two categories and eight subordinate concepts. First, the 'Information providing' type contains data, information or knowledge concerning a firm's activities, service usage, notice and review. Second, 'Advertisement' type suggests the price or brand name for users to purchase or use a product or service. Also this type contains the messages which make

Table 3. Examples of each type of information content

Variables		Definition	Keywords
Information providing	News	Reporting current activities about a firm	Current, partnership, interview
	Usage	Information on how to use service	Procedure, pay method, explain
	Review	Evaluating service or product improvements	Compare, review, postscript
	Notice	Introducing new products and giving announcements of a play, concert or special happening	Announce, launch, release
Advertisement	Sale	Suggesting price of products	Purchase, price
	Benefit	Suggesting use and benefit information about products	Discount, free, low price, compensation
	Event	Leading consumers to participate in events	Participation, event
	Public relations	Publicity concerning services or events	Event recommendations

users participate in events related to the product. Table 3 shows examples of each type of information content.

4.4 Analysis and Results

4.4.1 The Description of Information Contents

The intent was to examine whether olleh_mobile employed different tweets across the diffusion of tweets. This study considered two super-ordinate categories and eight sub-concepts. Of 886 tweets created by the olleh_mobile firm, “Information providing” tweets were about 64 % and “Advertisement” tweets were about 36 %. Results related to the four categories were 1) “Information providing” (News 5 %, Usage 28 %, Review 19 %, Notice 12 %) and the other four categories were in 2) “Advertisement” (Sale 2 %, Benefit 17 %, Event 13 %, PR 4 %). The results show that the olleh_mobile firm usually uses Twitter for informing customers about how to use products and services, and review of products and services.

Using the above characterization of tweet contents, this study determined whether differences in content appeared across the diffusion dimensions. Table 4 represents the results analysis of information diffusion according to information contents. The scale of RTs in “Information providing” type was news (6.52), usage (6.10), review (6.50) and notice (7.80). The speed of RTs in “Information providing” type was news (2.58), usage (19.20), review (16.20) and notice (15.60). The duration of RTs in “Information providing” type was news (30.50), usage (29.50), review (7.90) and notice (40.10). Also, the scale of RTs in “Advertisement” type was sale (4.80), benefit (18.40), event

Table 4. Information diffusion of information content

Constructs	Variables		Scale	Speed	Duration	N
Information content	Information providing	News	6.52	2.58	30.50	58
		Usage	6.10	19.20	29.50	336
		Review	6.50	16.20	7.90	225
		Notice	7.80	15.60	40.10	145
	Advertisement	Sale	4.80	1.20	13.20	29
		Benefit	18.40	3.00	62.10	199
		Event	21.50	1.02	74.40	159
		Service PR	7.10	1.56	31.98	46

(21.50), and service PR (7.10). The speed of RTs in “Advertisement” type was sale (1.20), benefit (3.00), event (1.02), and service PR (1.56). The duration of RTs in “Advertisement” type was sale (13.20), benefit (62.10), event (74.40), and service PR (31.98).

Findings indicated that Event and Benefit variables had the highest score occurring on the scale of RTs (21.50, 18.40) and duration of RTs (74.40, 62.10). The Usage variable had the highest score with respect to the speed of RTs. These results suggest that the advertisement type of tweets is more spread-out and the information type of tweets is shared most quickly. In addition, the result relating to the duration of RTs suggests that tweets or content do not last long. They remained an average of two days on Twitter.

4.5 The Impact of Information Contents on Diffusion

The first goal of this model was to evaluate whether there is a difference between information content related to information diffusion. Hypothesis H1 says that “The content types of information will have a significant effect on three types of information diffusion”. To examine this hypothesis, this study examined groupings across types of tweets by using cluster analysis which is typically utilized to examine patterns in various categories (Segars and Grover 1999).

This analysis separates data into groups that are represented by clusters, which can be meaningful, useful, or both. The clusters are constructed to be as internally homogenous as possible while also being as externally heterogeneous as possible. Numerous clustering algorithms have been used for the analysis of quantitative and qualitative data, and interested readers are encouraged to read the review of data clustering (Miaskiewicz and Monarchi 2008). Although several clustering algorithms exist, Ward’s minimum variance criterion was chosen for this analysis (Punj and Stewart 1983). The clustering criterion of this technique is minimization of total within-group sums of squares (Segars and Grover 1999).

The result of this analysis shows a three-cluster solution. Three clusters that emerged from the analysis were identified “Advertisement” (Benefit and Event),

Table 5. Description of information content clusters (means and standard deviation)

Variables	Cluster 1 (N = 189) “AD type”		Cluster 2 (N = 323) “IF type”		Cluster 3 (N = 374) “IF_AD type”	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Scale of RTs	19.28	29.43	4.01	5.06	5.96	6.56
Speed of RTs	1.16	2.45	19.70	171.14	15.76	125.32
Duration of RTs	66.50	133.87	27.02	97.37	13.80	53.43

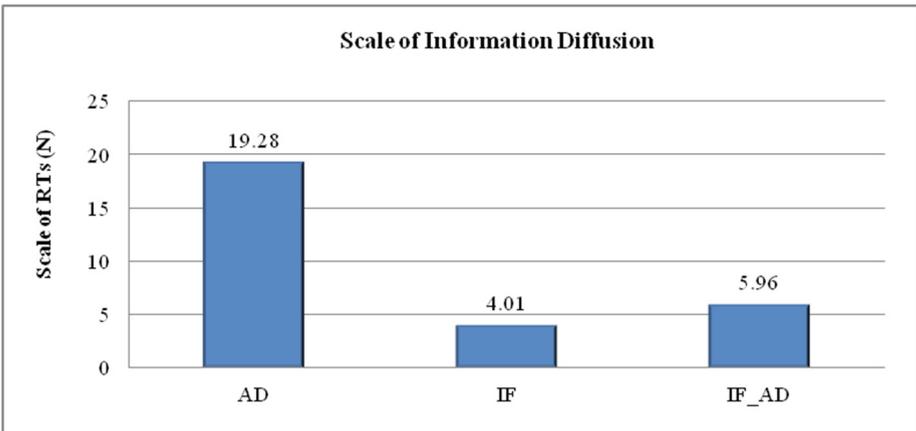


Fig. 2. The scale of information diffusion in three clusters

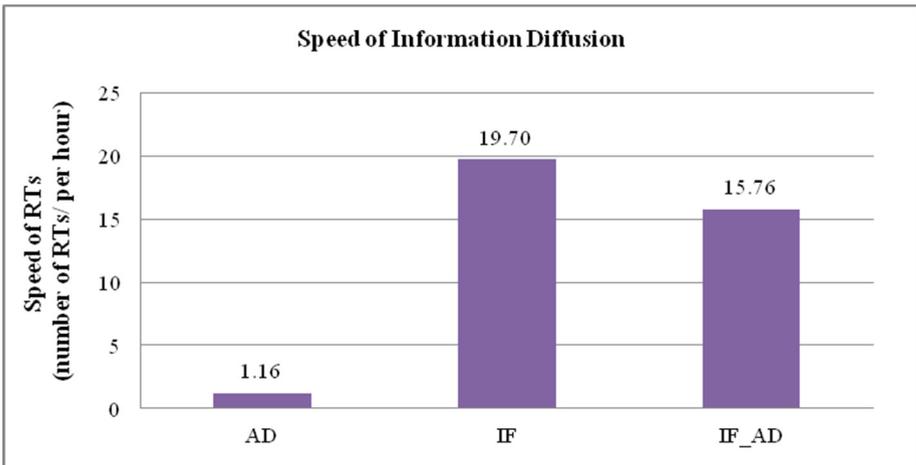


Fig. 3. The speed of information diffusion in three clusters

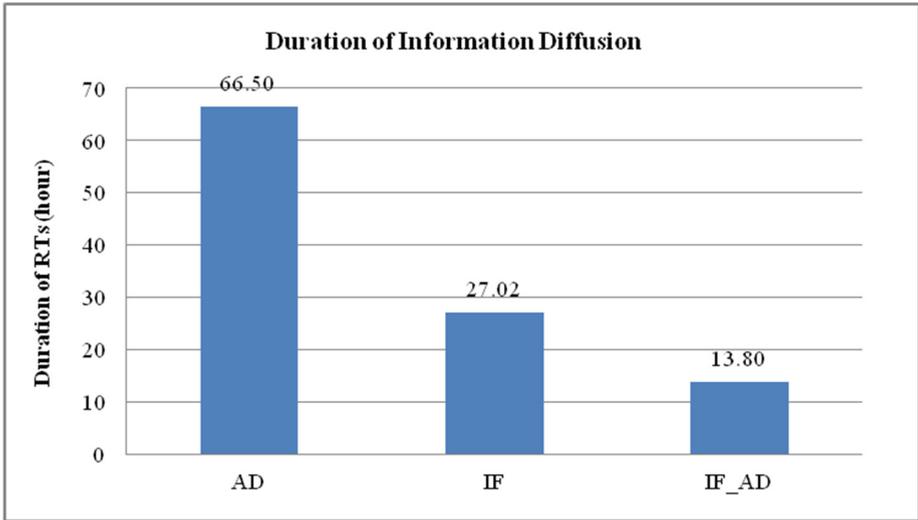


Fig. 4. The duration of information diffusion in three clusters

“Information providing” (Usage) and “Information and Advertisement” (Sale, Service PR, News, Review, Notice) as shown the Table 5.

The 189 tweets identified as “Advertisement” type are characterized by average scores on the scale of RTs, speed of RTs and duration of RTs. These results indicate that the scale ($M = 19.28$) and duration ($M = 66.50$) of RTs in this cluster are higher than other clusters. Also, the “Information” type cluster includes 323 tweets. This type has allow scale of RTs ($M = 4.01$). However, the speed of RTs ($M = 19.70$) is the highest score among three types of tweets. The cluster labeled “Information and Advertisement” consists of 374 tweets that show the lowest duration of RTs ($M = 13.80$). These results show that the “Advertisement” type is more associated with the scale and duration of RTs and “Information providing” type is more linked to the speed of RTs. Figures 2, 3 and 4 show the result of information diffusion according to information clusters.

After clusters of information contents were identified, a non-parametric test, the Kruskal-Wallis test was employed to explore the existence of diffusion differences. Within this study, differences between clusters are examined across the three dimensions of information diffusion. The Kruskal Wallis one-way analysis of variance by ranks does not assume a normal distribution, the analogous one-way analysis of variance. It is advantageous in statistical analysis to use ranks. The only assumptions underlying the use of ranks made are that the observations are all independent, that all those within a given sample come from a single population, and that the C populations are of approximately the same form (Kruskal and Wallis 1957). The test statistics are described by the following equation

$$W = \left[\frac{12}{N(N + 1)} \sum_{i=1}^k \frac{R_i^2}{n_i} \right] - 3(N + 1)$$

where k indicates the number of samples, and n_i and R_i mean the numbers of the observations and ranks in the i th sample, respectively.

Table 6 describes the summary statistics and correlation. Tables 7 and 8 suggest that the three-cluster solution represents meaningful differences in information content across the sampled firm. Table 8 outlines the results of the Kruskal-Wallis test for differences in information content across each dimension of information diffusion. These results indicate that the differences are significant for all three types of information content (Scale of RTs $p < .001$, Speed of RTs $p < .001$ and Duration of RTs $p < .001$). These findings provide support for hypothesis H1.

Table 6. Summary statistics and correlation matrix

	Mean	S.D	1	2	3	4	5	6
1. Scale of RTs	8.90	15.69	1.000					
2. Duration of RTs	29.86	97.45	.710**	1.000				
3. Speed of RTs	14.08	131.62	.511**	.129**	1.000			
4. IF_AD	1.09	1.45	-.218**	-.062	-.175**	1.000		
5. IF	0.84	0.99	.040	-.169**	.219**	-.647**	1.000	
6. AD	0.21	0.41	.208**	.277**	-.059	-.394**	-.445**	1.000

Table 7. Rank of Kruskal-Wallis test

	Cluster	N	Mean rank
Scale of RTs	Advertisement	189	545.10
	Information providing	374	455.56
	Information_Advertisement	323	370.09
	Total	886	
Duration of RTs	Advertisement	189	578.56
	Information providing	374	393.34
	Information_Advertisement	323	422.56
	Total	886	
Speed of RTs	Advertisement	189	415.51
	Information providing	374	508.58
	Information_Advertisement	323	384.52
	Total	886	

Table 8. Result of Kruskal-Wallis test

	Scale of RTs	Speed of RTs	Duration of RTs
Chi Square	57.769	44.173	70.054
DF	2	2	2
P value	.000	.000	.000

Note. **p < .05 ***p < .01

5 Conclusions

This study proposed an integrated research model based on diffusion of innovations and two-step flow theory. It classified information content and examined the effect of information content on diffusion of information. This study includes a number of theoretical implications. Previous studies focused on the characteristics of information within limited scopes (Rogers 1983; Rogers 2003). This study arrives at another perspective on information characteristics -information related to actual intentions or contents. Our results demonstrate that information content has a significant effect on information diffusion. When the type of information is “Information Providing”, information diffusion is fast. In the ‘Advertisement’ type, the scale of diffusion is higher and the duration is longer than for the ‘Information providing’ type. In other words, the patterns of information diffusion vary according to information content. For a given dataset of content created by a firm, Twitter information delivery to users occurs according to a wide variety of diffusion characteristics.

Our findings have important implications for business management. Our research provides useful guidelines in terms of understanding information content on Twitter. The knowledge that the ‘Information providing’ type typically associates with the speed of diffusion and the ‘Advertisement’ type associated with the scale and duration of diffusion provides business with helpful advice. It specifies not only how to quickly spread information but also how to make sure that users continue to repost tweets based on specific attributes of tweet contents. It shows that firms can spread information more quickly by providing the ‘Information and advertisement’ type rather than the ‘Advertisement’ type.

This study has performed a careful analysis of the impact of information contents and user characteristics on information diffusion. However, as with any empirical study, it has limitations. One issue arises from the tweets sample used and data collected because we choose a particular firm, olleh_Mobile. The impact of information contents for improving information diffusion are likely to differ among various firms. It is recommended that samples from a broad range of industries be used in future research. Thus, it is important to study how diffusion happens differently across specific industries.

Acknowledgement. This research was supported by the MSIP (Ministry of Science, ICT&-Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2014-H0301-14-1044) supervised by the NIPA (National ICT Industry Promotion Agency).

References

- Bajwa, D.S., Lewis, L.F., Pervan, G., Lai, V.S., Munkvold, B.E., Schwabe, G.: Factors in the global assimilation of collaborative information technologies: an exploratory investigation in five regions. *J. Manag. Inf. Syst.* **25**(1), 131–165 (2008)
- Barnes, S.J., Böhlinger, M.: Modeling use continuance behavior in microblogging services: the case of twitter. *J. Comput. Inf. Syst.* **51**(1), 1–10 (2011)
- Berinato, S.: Six ways to find value in twitter's noise. *Harvard Bus. Rev.* **88**(6), 34–35 (2010)
- Boettger, R.K., Palmer, L.A.: Quantitative content analysis: its use in technical communication. *IEEE Trans. Prof. Commun.* **53**(4), 346–357 (2010)
- Brancheau, J.C., Wetherbe, J.C.: The adoption of spreadsheet software: testing innovation diffusion theory in the context of end-user computing. *Inf. Syst. Res.* **1**(2), 115–143 (1990)
- Cha, M., Haddadi, H., Benevenuto, F., Gummadi, P.K.: Measuring user influence in twitter: The million follower fallacy. In: *Association for the Advancement of Artificial Intelligence* (2010)
- Chatman, E.A.: Diffusion theory: a review and test of a conceptual model in information diffusion. *J. Am. Soc. Inf. Sci.* **37**(6), 377–386 (1986)
- Cheung, C., M.K., Lee, M.K.O., Rabjohn, N.: The impact of electronic word-of-mouth. The adoption of online opinions in online customer communities. *Int. Res.* **18**(3), 229–247 (2008)
- Fichman, R.G.: Information technology diffusion: a review of empirical research. In: *Proceedings of the 13th International Conference on Information Systems* (1992)
- Fischer, E., Reuber, A.R.: Social interaction via new social media: (how) can interactions on twitter affect effectual thinking and behavior? *J. Bus. Ventur.* **26**(1), 1–18 (2011)
- Fullwood, C., Sheehan, N., Nicholls, W.: Blog function revisited: a content analysis of myspace blogs. *CyberPsychol. Behav.* **12**(6), 685–689 (2009)
- Greer, C.F., Ferguson, D.A.: Using twitter for promotion and branding: a content analysis of local television twitter sites. *J. Broadcast. Electron. Media.* **55**(2), 198–214 (2011)
- Ha, L., James, E.L.: Interactivity reexamined: a baseline analysis of early business web sites. *J. Broadcast. Electron. Media.* **42**, 457–474 (1998)
- Herring, S.C., Scheidt, L.A., Kouper, I., Wright, E.: *A Longitudinal Content Analysis of Weblogs: 2003-2004*. Routledge, London (2004)
- Jansen, B.J., Zhang, M., Sobel, K., Chowdury, A.: twitter power: tweets as electronic word of mouth. *J. Am. Soc. Inform. Sci. Technol.* **60**(11), 2169–2188 (2009)
- Java, A., Song, T.F., Tseng, B.: Why we twitter: understanding microblogging usage and communities. In: *Proceedings of the Ninth WEBKDD and First SNA-KDD Workshop on Web Mining and Social Network Analysis*, pp. 56–65 (2007)
- Kaplan, A.M., Haenlein, M.: Users of the world, unite! the challenges and opportunities of social media. *Bus. Horiz.* **53**(1), 59–68 (2010)
- Kim, H., Son, I., Lee, D.: The viral effect of online social network on new products promotion: investigating information diffusion on twitter. *J. Intell. Inf. Syst.* **18**(2) (2012)
- Kruskal, W.H., Wallis, W.A.: Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **47**(260), 583–621 (1957)
- Li, J., Rao, H.R.: Twitter as a rapido response news service: an exploration in the context of the 2008 china earthquake. *Electron. J. Inf. Syst. Dev. Countries.* **42**(4), 1–22 (2010)
- Lin, J.-S., Pena, J.: Are you following me? a content analysis of tv networks' brand communication on twitter' the purpose of using twitter. *J. Interact. Advertising.* **12**(1), 17–29 (2011)
- Miaskiewicz, T., Monarchi, D.E.: A review of the literature on the empathy construct using cluster analysis. *Commun. AIS* **22**, pp. 117–142 (2008)

- Morris, R.: Computerized content analysis in management research: a demonstration of advantages and limitations. *J. Manag.* **20**(4), 903–931 (1994)
- Naaman, M., Boase, J., Lai, C.-H.: Is it really about me? Message content in social awareness streams. In: *CSCW* (2010)
- Papacharissi, Z.: The blogger revolution? audiences as media producers. In: *The Annual Conference of the International Communication Association, New Orleans* (2004)
- Punj, G., Stewart, D.W.: Cluster analysis in marketing research: review and suggestions for application. *J. Mark. Res.* **20**, 134–148 (1983)
- Rogers, E.M.: *The Diffusion of Innovations*. Free Press, New York (1983)
- Rogers, E.M.: *Diffusion of Innovations*. Free Press, New York (2003)
- Rogers, E.M., Shoemaker, F.F.: *New York: Free Press* (1971)
- Savage, N.: Twitter as medium and message. *Commun. ACM* **54**(3), 18–20 (2011)
- Segars, A.H., Grover, V.: Profiles of strategic information systems planning. *Inf. Syst. Res.* **10**(3), 199 (1999)
- Simmons, L.L., Conlon, S., Mukhopadhyay, S., Yang, J.: A computer aided content analysis of online reviews. *J. Comput. Inf. Syst.* **52**, 43–55 (2011)
- Sledgianowski, D., Kulviwat, S.: Using social network sites: the effects of playfulness, critical mass and trust in a hedonic context. *J. Comput. Inf. Syst.* **49**(4), 74–83 (2009)
- Smith, B.G.: Socially distributing public relations: twitter, haiti, and interactivity in social media. *Pub. Relat. Rev.* **36**(4), 329–335 (2010)
- Thelwall, M., Buckley, K., Paltoglou, G.: Sentiment in twitter events. *J. Am. Soc. Inf. Sci. Technol.* **62**(2), 406–418 (2011)
- Yang, J., Counts, S.: Predicting the speed, scale, and range of information diffusion in twitter. In: *4th International AAAI Conference on Weblogs and Social Media (ICWSM)* (2010)
- Zhang, W., Watts, S.A.: Capitalizing on content: information adoption in two online communities. *J. Assoc. Inf. Syst.* **9**(2), 72–93 (2008)
- Zhao, D., Rosson, M.B.: How and why people twitter: the role that micoblogging plays in informal communication at work. In: *Proceedings of the ACM 2009 International Conference on Supporting Group Work, Sanibel Island, FL*, pp. 243–252 (2009)



<http://www.springer.com/978-3-319-16867-8>

Scalable Information Systems

5th International Conference, INFOSCALE 2014, Seoul,
South Korea, September 25-26, 2014, Revised Selected
Papers

Jung, J.J.; Badica, C.; Kiss, A. (Eds.)

2015, IX, 107 p. 33 illus., Softcover

ISBN: 978-3-319-16867-8