

Analysis of Social Influence and Information Dissemination in Social Media: The Case of Twitter

Chien-Wen Shen and Chin-Jin Kuo

Department of Business Administration, National Central University
No.300, Jhongda Rd., Jhongli City 32001, Taiwan
cwshen@ncu.edu.tw

Abstract. To understand how the influencers of event information dissemination on social media can be identified, we propose three perspectives for investigating this topic: the number of related messages posted by the influencers, the number of related messages in which the influencers are mentioned by other users, and the number of influencers' messages that are reposted. The findings regarding social influencers can help companies identify the key people or organizations with whom they must engage. In addition, we used a social network diagram to depict how event information is disseminated from the influencers. This diagram shows the top influencers at different stages of information dissemination. Effectively modeling relationships among top users and accordingly using them to filter or recommend information are fundamental for mining social networking services. To illustrate our approach, we used the tweets from the Windows 8.1 launch event as a case study.

Keywords: Social Influence, Information Dissemination, Social Networks, Twitter.

1 Introduction

In an era of social media, people do not merely rely on traditional media that delivers information through one-way communication channels. People directly engage and become “the media” by using various online platforms such as content communities, social networking sites, blogs and microblogs to share their opinions, insights, information, experiences, and perspectives with each other [1]. On social networking services such as Facebook and Twitter, users build their own friendship networks and widely share, discover, and spread information by using various formats such as words, pictures, audio, and video. Because of the extraordinary popularity of social media, more than 10 billion messages are currently on Facebook daily, and more than 300 billion messages have been sent on Twitter [2]. Hence, understanding the characteristics of information dissemination in social media has become a crucial topic both in academia and business. Related research from Cha et al. [3] investigated the 11 million photos shared by 2.5 million users on Flickr to determine how widely and quickly the information spread through the social network as well as regional influences on popularity. Kwak, Lee, Park, and

Moon [4] collected 41.7 million user profiles and 106 million tweets to study the topological characteristics of Twitter and its power as a new medium of information sharing from the perspectives of connectivity, influencers, and retweets. Ediger et al. [5] used data sets comprising H1N1 tweets and Atlanta Flood tweets from 2009 to identify the top vertices of tree-structured message connections in Twitter. Their findings indicated that major media outlets and government organizations are usually the highest-degree vertices in news dissemination. Aral and Walker [6] designed and conducted a randomized field experiment that evaluate the effectiveness of active-personalized referrals and passive-broadcast notifications in creating peer influence and social contagion among the 1.4 million friends of 9,687 experimental users on Facebook. Yoganasimhan [7] examined a dataset of 1939 YouTube videos uploaded in 2007 to understand how the size and structure of a user's local network affect its ability to disseminate information on YouTube. Lin, Lazer, and Cao [8] introduced a visualization tool that highlights the social and spatiotemporal processes of information diffusion through social media in real time. This tool was designed based on a metaphorical sunflower of which the seeds are often dispersed far away.

To provide further insight into information sharing using social media, one of the objectives of this study is to discuss how the influencers of event information dissemination on social media can be identified. We propose three perspectives for investigating this topic: the number of related messages posted by the influencers, the number of related messages in which the influencers are mentioned by other users, and the number of influencers' messages that are reposted. Because the perceptions of brands and companies are largely driven by the rapid expansion of social media channels through which influencers communicate with their social networks [9], findings regarding influencers can help companies identify the key people or organizations with whom they must engage. In addition, we used a social network diagram to depict how event information is disseminated from the influencers. This diagram shows the top influencers at different stages of information dissemination. Analyzing the process of information dissemination on social media is complex because a numerous people use social media and their complex topological relationships must be considered. Effectively modeling relationships among top users and accordingly using them to filter or recommend information are fundamental for mining social networking services. The methodology used for the aforementioned objectives is detailed in the next section. To illustrate our approach, we used the tweets from the Windows 8.1 launch event as a case study, which is discussed in Section 3. Finally, Section 4 concludes our study.

2 Methodology

To analyze the characteristics of an event information dissemination on social media, the related key words must first be identified, and the related posts are then retrieved by using the tools provided by social networking services. For example, on Twitter, the streaming application programming interface (API) can be used to stream tweets in real time, or the search API can be used to perform ad hoc user queries from a limited corpus of recent tweets. Based on the collected data set, the most intuitive method for identifying

the influencers of event information dissemination on social media is to count the number of related messages posted by users. Because different social networking services have distinct data structures or metadata tags, the related data fields must be understood. For example, on Twitter, the user who posted a message can be identified by referring to the `user_screen_name` data field. Hence, the users who have the highest frequency numbers in `user_screen_name` among the tweets related to the target event in a given observation time can be considered the influencers of information dissemination. In this study, we selected only the top 10 influencers to avoid generating excessively complex discussions. Moreover, a user mention in social media is any message update that contains a person's username in the body of the message. Hence, the users who receive the highest frequencies of mentions in the target social media feeds can also be considered influencers of information dissemination. On Twitter, user mention data can be retrieved by a code like the follows.

```
if doc['entities'].get('user_mentions'):
    for user_mention in
doc['entities']['user_mentions']:
    yield('@' + user_mention
['screen_name'].lower(), [doc['_id'], doc['id']])
```

Because the value of the attribute `screen_name` in `user_mentions` indicates the username mentioned in a message, the users with the highest counts of `screen_name` in the target tweets can be considered the influencers of the related information disseminated on Twitter. In addition, the users who frequently repost information related to an event were also considered influencers. On Twitter, reposting of another person's tweet is called a retweet. This feature allows a user's retweets to be disseminated to all of the user's followers. The number of times that a username appears after the metadata tags "RT" or "via" in the target tweets can be counted to determine how many related tweets user has retweeted. After identifying the influencers, we further investigated how information is disseminated from these influencers by depicting a social network diagram. A node representing an influencer is allocated in the center of a network diagram and is associated with the top five users who have reposted the most messages from this influencer. This procedure can be repeated to obtain a social network diagram that illustrates how event information is disseminated from influencers as well as the influencers of information dissemination at different stages. The influencers at the first and the second stages are essential for the initial distribution of event information, and the influencers at the remaining stages are responsible for the subsequent information circulation.

3 Case Study

In this study, we used Twitter as the form of social media example to illustrate our approach. Twitter is the most popular microblogging platform, and its prominence as a social media platform is growing [10]. According to a report conducted by the Pew Research Center [11], 72% of U.S. adults online use social networking sites and 18% of them are Twitter users. Because of its great popularity, Twitter has

been exploited as a platform for the viral marketing of content, products, and political campaigns [12]. In this case study, we used the Windows 8.1 product launch event to demonstrate how the information from this event was disseminated on Twitter. Microsoft launched Windows 8.1 on October 17, 2013, and a 1-month collection of tweets that included the key word “Windows 8.1” was obtained by using the Twitter search API. A total of 450,250 messages published by 206,792 users were collected. Fig. 1 shows that nearly 60% of the related tweets were published during the first week after the Windows 8.1 launch date. The daily tweets drastically dropped from more than 100,000 to less than 20,000 within 4 days. However, the number of tweets increased to more than 20,000 on October 21 because Microsoft temporarily pulled the Windows RT 8.1 update from the Windows Store. Since that date, the tweets regarding Windows 8.1 continued decreasing. One month after the Windows 8.1 launch date, the number of related tweets was only approximately 2,000. Hence, one month of data collection is sufficiently to effectively investigate the issues of social influence and information dissemination regarding Windows 8.1 launch event on Twitter.

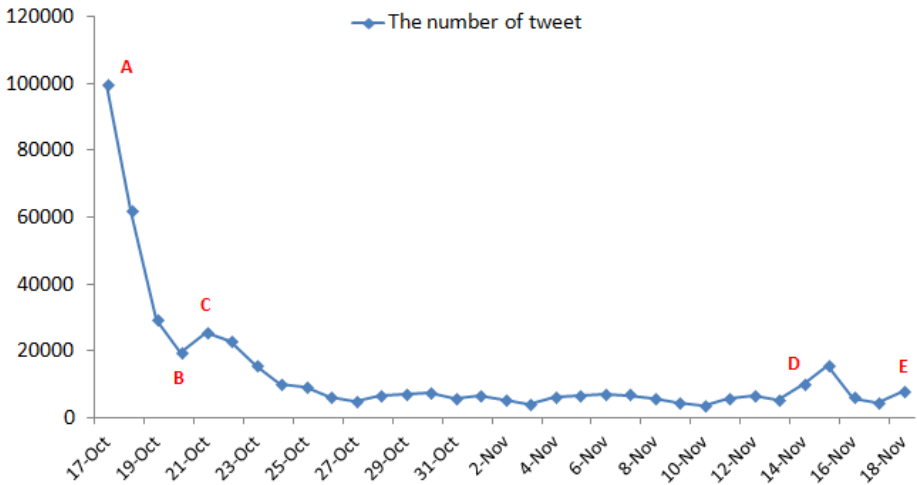


Fig. 1. Distribution of Windows 8.1 tweets during the first month after the launch date

To identify the social influencers of information dissemination regarding Windows 8.1 event on Twitter, we commenced with an investigation of the top users who posted the greatest number of related tweets. According to the results shown in Table 1, the top 10 users (@Windows, @hatiwin, @WindowsSupport, @atlazone, @arsenico, @NotebookDeal, @softpediadriverr, @WinObs, @jorin5712, @scoop_india) posted an average of 723 related tweets during the first month following the launch of Windows 8.1. However, the average number of tweets related to Windows 8.1 posted by general users was only 2.18. Hence, these influencers shared 100 times more related tweets than the average users did. Among the top 10 users, only the users @Windows and @WindowsSupport are from the official Microsoft accounts. Besides, the total tweets of the top 10 users account for only 1.60% of the investigated Windows 8.1

tweets posted by 206,792 Twitter users. This number indicates that numerous people were engaged in sharing information regarding the Windows 8.1 launch event on Twitter, and that this information dissemination was not limited to a small group of users.

Table 1. Top 10 influencers by the number of tweets

Rank	Username	Number of Tweets	Percentage
1	@Windows	1,110	0.25%
2	@hatiwin	1,102	0.24%
3	@WindowsSupport	1,037	0.23%
4	@atlazone	775	0.17%
5	@arsenico	708	0.16%
6	@NotebookDeal	683	0.15%
7	@softpediadriver	669	0.15%
8	@WinObs	427	0.09%
9	@jorin5712	361	0.08%
10	@scoop_india	359	0.08%
Total		7,231	1.60%

In addition, we identified the social influencers of information dissemination on Twitter by investigating the number of related Windows 8.1 messages in which other users mentioned the influencers. The mention mechanism on Twitter is often used to establish conversations between users through the exchange of messages or simply by referring to a person in the message's text [13]. Table 2 shows a summary of the top 10 influencers of the Windows 8.1 launch event based on the number of mentions. Unsurprisingly, Microsoft's official accounts @Windows and @Microsoft were among the most mentioned usernames in the related tweets. Around 13.21% of the related tweets were disseminated from these Microsoft's official accounts. Although @ishibasystems, @youtube, @detiknews5, @sambung_cerita, @henextweb, @artem_klyushin, @mashable, and @Windowsblog are not official Microsoft company accounts, they were also mentioned by numerous Twitter users. These non-official Twitter users accounted for 16.65% of all mentions; therefore, Microsoft may consider establishing close relationships with these influencers because they were crucial sources of information regarding the Windows 8.1 launch event for Twitter users.

Another key mechanism Twitter provides is the retweet feature, which enables people to quickly share tweets with their followers. This feature enables individual tweets to be propagated throughout social networks and serves as a method for people to endorse their perspectives regarding specific topics [13]. Therefore, the influencers of information dissemination on Twitter can be identified according to the number of retweets by other users. Among the 450,250 related tweets during the first month since the Windows 8.1 launch date, 19.92% were retweets. Table 3 lists the top 10 users who had the greatest number of messages reposted by other users. Because the retweets from these influencers accounted for 26.87% of total retweets, these influencers were vital to the Windows 8.1 event information disseminated on Twitter.

In addition to the Windows and Microsoft official accounts, most of the influencers were non-official users. Hence, Microsoft must quickly update these users with current information because many of their followers spread news regarding Windows 8.1 provided by these influencers.

Table 2. Top 10 influencers by the number of mentions

Rank	User ID	Count	Percentage
1	@Windows	14,362	10.06%
2	@ishibasystems	4,967	3.48%
3	@Microsoft	4,505	3.15%
4	@youtube	4,125	2.89%
5	@detiknews5	3,599	2.52%
6	@sambung_cerita	3,216	2.25%
7	@thenextweb	2,437	1.71%
8	@artem_klyushin	2,034	1.42%
9	@mashable	1,777	1.24%
10	@Windowsblog	1,615	1.13%
Total		42,637	29.86%

Table 3. Top 10 influencers by the number of retweets

Rank	User ID	Count	Percentage
1	@ishibasystems	4,597	5.13%
2	@Windows	4,137	4.61%
3	@detiknews5	3,595	4.01%
4	@thenextweb	2,365	2.64%
5	@artem_klyushin	2,034	2.27%
6	@Microsoft	1,685	1.88%
7	@sambung_cerita	1,607	1.79%
8	@mashable	1,380	1.54%
9	@cnet	1,371	1.53%
10	@kakocom	1,327	1.48%
Total		24,098	26.87%

A social network diagram illustrating how the Windows 8.1 tweets were disseminated from the official Windows and Microsoft Twitter accounts to the other users is depicted in Fig. 2. In this diagram, the Windows and Microsoft nodes are centrally located and are tied to the nodes that represent the top users (followers) who have reposted the most Windows 8.1-related messages from the official Windows and Microsoft accounts. Fig. 2 shows that the top followers of the Windows account were @microsoft_now, @windowsphone, @MicrosoftID, and @WinObs. The top followers of the Microsoft account were @MicrosoftASIA, @dovellonsky, @MicrosoftTH, @BenThePCGuy, and @lee_stott. In addition, the top followers that resent the tweets from the top followers of @Windows and @Microsoft were also illustrated in Fig. 2.

Hence, a total of 3 layers of top influencers related to the information dissemination of Windows 8.1 launch event were identified.

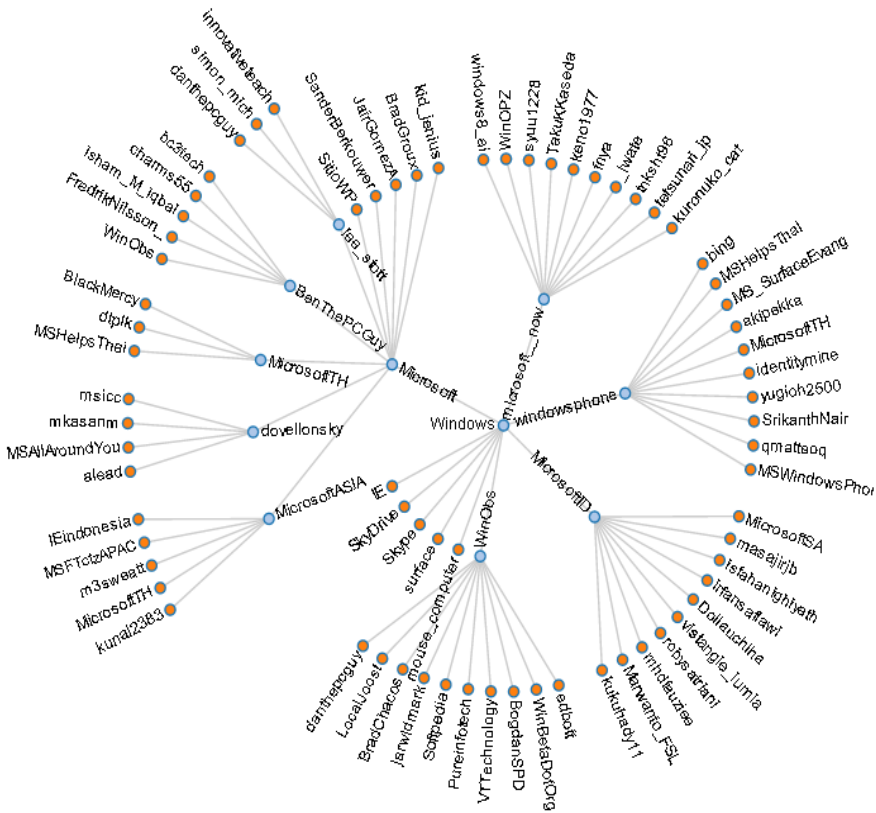


Fig. 2. A social network diagram that identifies the top influencers related to the information dissemination of Windows 8.1 launch event

In addition, we measured the strength of the ties between a social influencer and their top followers according to the number of the user’s tweets that were reposted by followers. Table 4 shows the results regarding the strength of the tie between the nodes in Fig. 2. The results revealed that 4,137 retweets were from the official Windows account, and that 1,647 retweets were from the official Microsoft account. Users who exhibit strong retweet influence are highly capable of generating content with pass-along value. The influences at the two initial stages were essential for the initial distribution of Windows 8.1 event information. The social influencers at the later stages were critical for the dissemination of this information on Twitter.

Table 4. Strength of the ties between social influencers and their followers

Stage-1 Influencer	Stage-2 Influencers	Number of Retweets	Stage-3 Influencers	Number of Retweets
@Windows (4137)	@Microsoft	1647	@MicrosoftASIA	29
			@dovellonsky	21
			@MicrosoftTH	13
			@BenThePCGuy	13
			@lee_stott	12
	@microsoft_now	334	@windows8_ai	52
			@WinOPZ	33
			@syuu1228	23
			@TakuKKaseda	7
			@keno1977	4
	@windowsphone	305	@bing	45
			@MSHelpsThai	37
			@MS_SurfaceEvang	34
			@akipekka	34
			@MicrosoftTH	13
@MicrosoftID	268	@MicrosoftSA	25	
		@masajirjb	21	
		@isfahanighiyath	7	
		@irfansaflawi	4	
		@Dollauchiha	4	
@WinObs	252	@edbott	207	
		@WinBetaDotOrg	155	
		@BogdanSPD	135	
		@VTTechnology	132	
		@Pureinfotech	122	

4 Conclusion

Because of the viral nature of information dissemination on social media, identifying and analyzing the influential social media users is becoming increasingly important to business. This study not only provides different perspectives for investigating the influencers of event information dissemination, but also proposes a social network diagram approach to depict how event information is disseminated from the influencers. Our proposed approaches can help companies identify the key users during an event information dissemination as well as model the relationships among top influencers. A case study regarding the tweets from the Windows 8.1 launch event was also discussed to illustrate our approaches. Our results suggest that these social influencers did play important roles on the information dissemination of the Windows 8.1 launch event. We are also able to understand how event information was disseminated from the influencers based on the social network diagram. Future studies may

explore other events to evaluate our proposed approaches on identifying the influencers. Besides, we can investigate how our approach can be applied on other social networking platforms such as Facebook and YouTube. We can also further examine how public sentiment towards an event can be assessed based on the related messages on social media.

Acknowledgements. This research was supported in part by the Ministry of Science and Technology, Taiwan (grant # MOST 103-2410-H-008 -029).

References

1. Thevenot, G.: Blogging as a Social Media. *Tourism and Hospitality Research* 7, 287–289 (2007)
2. Smith, C.: By the Numbers: 130 Amazing Facebook User & Demographic Statistics, <http://expandedramblings.com/index.php/by-the-numbers-17-amazing-facebook-stats/>
3. Cha, M., Mislove, A., Gummadi, K.P.: A Measurement-Driven Analysis of Information Propagation in the Flickr Social Network. In: *Proceedings of the 18th International Conference on World Wide Web*, pp. 721–730. ACM, Madrid (2009)
4. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter, a Social Network or a News Media? In: *Proceedings of the 19th International Conference on World Wide Web*, pp. 591–600. ACM, Raleigh (2010)
5. Ediger, D., Jiang, K., Riedy, J., Bader, D.A., Corley, C., Farber, R., Reynolds, W.N.: Massive Social Network Analysis: Mining Twitter for Social Good. In: *2010 39th International Conference on Parallel Processing*, pp. 583–593 (2010)
6. Aral, S., Walker, D.: Creating Social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks. *Management Science* 57, 1623–1639 (2011)
7. Yoganarasimhan, H.: Impact of Social Network Structure on Content Propagation: A Study Using YouTube Data. *Quantitative Marketing and Economics* 10, 111–150 (2012)
8. Lin, Y.-R., Lazer, D., Cao, N.: Watching How Ideas Spread Over Social Media. *Leonardo* 46, 277–277 (2013)
9. Booth, N., Matic, J.A.: Mapping and Leveraging Influencers in Social Media to Shape Corporate Brand Perceptions. *Corporate Communications* 16, 184–191 (2011)
10. Efron, M.: Information Search and Retrieval in Microblogs. *Journal of the American Society for Information Science and Technology* 62, 996–1008 (2011)
11. Brenner, J., Smith, A.: 72% of Online Adults are Social Networking Site Users. Pew Research Center's Internet & American Life Project, Pew Research Center (2013)
12. Jin, F., Dougherty, E., Saraf, P., Cao, Y., Ramakrishnan, N.: Epidemiological Modeling of News and Rumors on Twitter. In: *Proceedings of the 7th Workshop on Social Network Mining and Analysis*, pp. 1–9 (2013)
13. Borondo, J., Morales, A., Losada, J., Benito, R.: Characterizing and Modeling an Electoral Campaign in the Context of Twitter: 2011 Spanish Presidential Election as a Case Study. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 22, 23138 (2012)