10-8-2018

Product Popularity versus Size of Conversation in Social Media: An Analysis of Twitter Conversations about YouTube Product Categories

Nitin Mayande
Portland State University

Charles Weber
Portland State University, webercm@pdx.edu

Citation Details

N. Mayande and C. Weber, "Product Popularity versus Size of Conversation in Social Media: An Analysis of Twitter Conversations about YouTube Product Categories," 2018 Portland International Conference on Management of Engineering and Technology (PICMET), Honolulu, HI, 2018, pp. 1-7.

Product Popularity versus Size of Conversation in Social Media: An Analysis of Twitter Conversations about YouTube Product Categories

Nitin Mayande, Charles Weber
Dept. of Engineering and Technology Management, Portland State University, Portland, Oregon, USA

***Abstract*—It has always been assumed that a large conversation about a topic on social media implies that the topic is popular. However, an empirical study of Twitter conversations about a variety of YouTube product categories, which is described in this paper, has shown that this is not necessarily the case. Popularity as measured by distribution volume is not necessarily a reliable indicator of the size of a community or conversation that is associated with a product category. This suggests that current online marketing practices are not nearly as effective as has been assumed to date. Novel, potentially more effective approaches to online marketing are suggested in the paper.**

I. INTRODUCTION

People all around the world are utilizing online social networks at an astonishing rate, and today's marketers are responding to the increasing importance of online social networks by spending billions of dollars in digital marketing *(Ng and Vranka [42])*. With increased spending on social media, businesses are feeling the pressure to gain new insights into customer behavior *(Halavais [28], Lindsay, et al., [35])*. Success in marketing through online social media apparently critically depends upon understanding the social network that may have a potential interest in your product or service and by identifying the key attributes about the influencers that will spread your marketing message *(Lindsay, et al. [35])*. Yet, this is easier said than done, because to date nobody really understands how online social networks get organized *(Aral, et al. [3], Cha, et al. [12] [13], Li and Bernoff [33], Mayande [37], Weber and Mayande [52], Wiertz, et al. [54])*.

This lack of understanding can have severe consequences such as gross misallocation of marketing resources *(Edwards [18] [19])* and an inability to shape online conversations about a product *(Chakrabarti and Berthon [14])*. The latter can result in missed opportunities such as free advertising and better brand recognition *(Longart [36])*. Even worse, negative conversations about a product that can lead to irreparable financial damage to the firm that develops the product *(Ayres [6], Khammash and Griffiths [33])*.

An issue of preeminent interest in online marketing is whether activity transcends platforms. For example, does the number of online conversations on a social networking platform about a specific product category on another media platform indicate how popular that product category is? Is the size of the online community that discusses the product category an indicator of the product's popularity?

This paper describes an empirical study that tests the following guiding proposition: *The popularity of a product category is reflected in the size of the community and the magnitude of the online conversations that discuss the product category*. In other words, more popular product categories should generate communities that are bigger in size. These communities, in turn, should generate larger conversations about the product category. Conversely, accessing larger conversations should consequently grant marketers access to more popular products.

To test the abovementioned proposition, we, the authors of this paper, analyze the metadata of online conversations about six product categories that vary greatly in popularity, where distribution volume acts as a proxy for popularity. Establishing a positive correlation between the distribution volume of a product category and the size and activity level of the online community that discusses the product category would support the above proposition. A failure to do so would suggest that activity does not transcend platforms. In that case, engaging blindly with the largest online community that discusses a specific product category on another platform may constitute a misallocation of resources, because that product category may not be popular.

II. CASE STUDY RESEARCH METHOD

Fig. 1 illustrates the design of this study, which deploys the case study research method *(Eisenhardt [20], Yin [57])* to look at social networks from the point of view of product categories. A product category that is discussed by a social network is considered a case. The product category in each case is sufficiently mature, so as to avoid any bias associated with startup effects. Conversely, the product category should not be in rapid decline, so as to avoid any bias that pertains to rapid decay of the social network under study.

978-1-890843-37-3 ©2018 PICMET

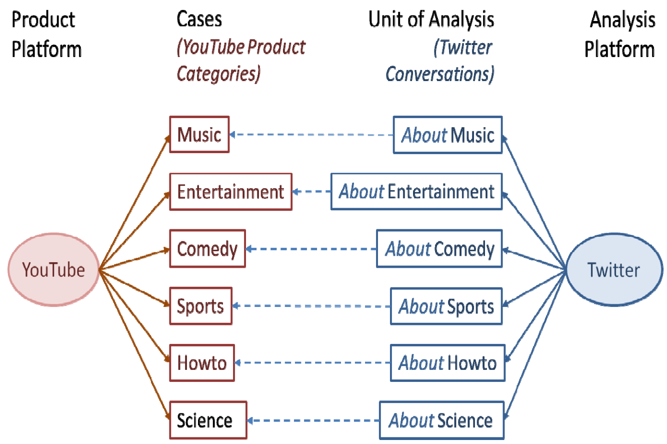

Figure 1: Research Design

A. Case Selection

Case selection in this study (like in many others) depends upon theoretical sampling and replication logic *(Leonard-Barton [32], Yin [57])*. The key criterion for theoretical sampling in this study is scale, because we would like to find out whether the social networks that discuss products categories in which content is consumed at high volumes behave differently from social networks that discuss product categories in which content is consumed at relatively low volumes. Replication logic manifests itself by selecting two product categories from each level of distribution volume. Thus the guiding proposition of this research is tested in more than one case. However, replication of cases "requires that the phenomenon being studied be defined by some characteristics common to all the research situations" *(Yin, [55], as cited by Leonard-Barton [32], p. 251)*. Thus, all cases in this research come from a common delivery platform—YouTube.

B. Delivery Platform: YouTube

The success of a product category delivered on YouTube depends on its "popularity" or distribution volume, which is generally measured by the total number of views per unit time *(Xu, et al. [56])*. Theoretical sampling *(Eisenhardt [20], Leonard-Barton [32], Yin [57])* in this study consequently consists of choosing product classes that either have very high or relatively low distribution volumes, as well as some product classes of intermediate scale. Each YouTube product class consequently constitutes a case. Replication logic dictates that our sample should contain at least two of each, *i.e.* a total of at least six.

The YouTube delivery platform was chosen as a setting for this research because of its wide variability in scale. Some YouTube product categories are an order of magnitude more popular than others. "Music," "Comedy," "Entertainment," and "Sports" have been identified as categories of interest on YouTube in the academic literature *(Thelwall, et al. [49], Xu, et al. [56])*, as well as in industry reports *(Sysomos [47])*. "Music" has been rated as the most popular category, as it comprises of almost 31% of all videos. "Entertainment" has been slated to be the second most popular category with 14.59% of all videos. Music and Entertainment have consequently been chosen as cases in the "large" volume category. "Comedy" and "Sports" categories are in the middle range of popularity with each category comprising of almost 6% of all videos. They will serve as cases in the "medium" category. We also intend to analyze "Howto" and "Science" categories, as they lie on the lower end of popularity, comparatively, with each category comprising only 3% and 2.5% of overall videos, respectively.

C. Analysis Platform: Twitter

Twitter was chosen as the analysis platform for this study for the following reasons. First, Twitter is the only social media platform that can capture changes in the context and content of online conversations at the rate at which they actually occur. Furthermore, all data on Twitter are available in the public domain. Finally, Twitter is popular enough for it to cover a sufficient number of conversations to enable a comprehensive analysis of the product categories under study. As early as 2014, Twitter received almost 190 million unique visits every month *(Alexa [2])*, which makes it the eighth most popular website in the world, and over 1 billion tweets were generated on Twitter every 5 days *(Statisticbrain [46])*.

Twitter is a micro-blogging platform *(Zhao and Rosson [58])*, which was founded in 2006. Microblogs are short comments usually delivered to a network of associates *(Huberman, et al. [30])*. Microblogging is also referred to as micro-sharing, micro-updating, or 'tweeting' *(Huberman, et al. [30])*. A message on Twitter is known as a 'tweet'.

Every person or entity on Twitter (like alias, company, etc.) is identified by its Twitter handle. Every Twitter handle can tweet. A Twitter handle can direct a tweet towards another Twitter handle by "@ mentioning" them. The recipient Twitter handle can either forward the message to its network by retweeting "RT @" the sender's message, or reply to the sender by "@ mentioning" the sender's Twitter handle. The recipient can choose to do neither.

Tweets have a very unique character. In contrast to many other messages, they are limited to 140 characters *(Ramage, et al. [45])*. Tweets commonly ask for or share information, news, opinions, complaints, or details about daily activities. Tweets may include hyperlinks to news stories, blogs, pictures, videos, etc. Tweets show up in the stream of those following the poster of the tweet; most posts are also publically available. Tweets are time stamped and publicly displayed on the Twitter platform.

Tweeting directly impacts word of mouth communication because it allows people to share thoughts almost anywhere (i.e., while driving, getting coffee, or sitting at their computer) to almost anyone "connected" (e.g. Web, cell phone, instant messaging, email) on a scale that has not been seen in the past *(Honeycutt and Herring [29])*. While the shortness of the microblog keeps people from writing long thoughts, it is precisely the micro part that differentiates microblogs from other word-of-mouth media, including full blogs, web pages, and online reviews *(Ramage, et al. [45])*. A standard microblog is approximately the length of a typical newspaper headline and subheading *(Milstein, et al. [41])*, which makes it easy to both produce and consume.

D. Population Study

The study described in this paper is a population study. Modern data extraction capabilities on the Internet allow us to study whole populations. This approach not only eliminates sample selection bias; it also ensures that the results observed are valid and generalizable to the entire population under study. This is especially important in studies that involve networks, as selecting only a sample instead of the population can break a network into multiple small networks *(Goggins and Petakovic [26])*, which can lead to faulty results. Furthermore, the data collection method deployed in this study (see section F) allows us to extract large amount of data from which statistically significant conclusions can be drawn.

E. Data Collection

We have conducted a retrospective study, for which data were collected in continuous time. Under these circumstances, the number and sequence of events and the duration between them can all be calculated. The main advantage of this approach lies in the greater detail and precision of information *(Blossfeld and Rohwer [8])*. It also reduces time required to collect data, and it enhances the chances of recognizing the overall patterns *(Leonard-Barton [32])*.

Given that YouTube is the platform of analysis, each YouTube product category constitutes a case. Twitter is the research setting, and Twitter conversations about specific YouTube product categories have become the unit of analysis. Data on the conversations about the chosen product categories were collected on Twitter.

Twitter data is easily available through application programming interfaces (API's) from which the networks forming within a context can be easily deduced. For the sake of simplicity, we use keyword search as a means of finding contextual networks *(Jansen, et al. [31])*. Both the Twitter platform as a data source and keyword search as data filter have been used in previous studies *(Jansen, et al. [31], Teevan, et al. [48], Williams, et al. [55])*.

Data were gathered for a period of three months (a total of 91 days), from December 31st, 2013 to March 31st, 2014, in order to control for any monthly periodicity in the data *(Gonçalves and Ramasco [27], Meiss, et al. [40])*. The particular time period of data collection was chosen at random. The data have been analyzed in daily intervals, in order to capture tweet volatility patterns caused by daily routine *(Dodds et al., [17])*. (For example, Twitter users in Tokyo tweet a lot less during working hours *(Gigaom [25])*.) The 24 hours started in accordance with Greenwich Meridian Time (GMT).

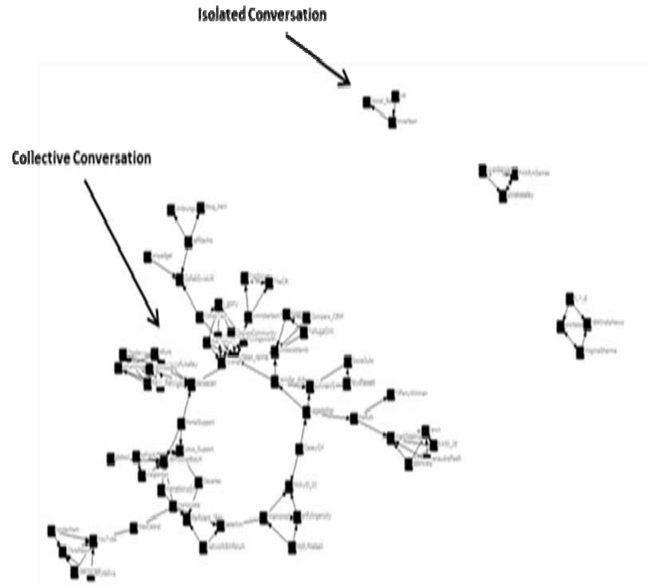

Figure 1: Collective vs. Isolated Conversations

F. Managing Noise

Twitter generates more than 1 billion tweets every 5 days *(Statisticbrain [46])*. Therefore, in order to reach the relevant audience, it is important to weed out noise, which is classified into two categories:

1. *Contextual Noise:* People have multiple topics of interest which may vary from the work that they do, their hobbies, their likes and dislikes, lifestyle choices, etc. Hence, they tweet about these multiple topics of interest. In order to identify a relevant social network, the context of conversations that is relevant to the business objectives (marketing, brand perception, customer support, etc.) needs to be identified. The remaining conversations fall under contextual noise. Contextual noise is very subjective and depends upon the business objective. Reducing contextual noise is achieved by using keyword searches.
2. *Broadcast Noise:* After identifying the context, a social network forming within that context can be identified. In order to identify these networks, it is necessary to identify the relationships people form within the network. Relationships in this case are formed when people interact with each other. In this case, we consider two actions that form relationships when they are tweeting somebody: @ mentioning or retweeting (RT @). The tweets that do not evoke any response, i.e., nobody interacts (@mentions or RT @), are considered broadcast noise. The rate of participation in the largest network does not impact the size of network, but it does impact the volume of tweets associated with the largest network. Therefore, while considering the total number of people participating in the largest network, only the Twitter user names that participate on a particular day are counted for that day. Even if the participants tweet more than once, they are still only counted once as the 'daily unique'. But, while

considering the total number of tweets, only the tweets associated with the largest network will be counted for analysis. The same process will be followed while measuring number of people participating on daily basis and tweet volumes on a daily basis associated with overall topic, broadcast and engaged activity within the overall topic.

The removal of broadcast noise allows us to analyze people that are engaged in the contextual conversation. Within that conversation only the largest network of people (community) that is engaged in a collective conversation everyday will be considered for analysis. The distinction between the collective conversation and isolated conversations is shown in Figure 2. A large group of people are engaged in a collective conversation, whereas small isolated groups converse on the side in isolated conversations.

TABLE 1: DEFINITIONS OF METADATA

Total_Tweets: Cumulative sum of daily volume of tweets associated with the topic.

Broadcast_Tweets: Cumulative sum of daily volume of tweets associated with all the broadcast activity on a topic. These are all the tweets without an @mention.

Engaged_Tweets: Cumulative sum of daily volume of tweets associated with all the engaged activity in a topic. These are the tweets with an @mention.

Community_Tweets: Cumulative sum of daily volume of tweets associated with the largest network engaged in the collective conversations within a topic (see fig. 2).

Total_People: Cumulative sum of daily unique people associated with the topic.

Broadcast_People: Cumulative sum of daily unique people associated with all broadcast activity in a topic.

Engaged_People: Cumulative sum of daily unique people associated with all engaged activity in a topic.

Largest_Community: Cumulative sum of daily unique people associated with the largest network engaged in the collective conversation within a topic.

G. Variables and Measures

Scale becomes a control variable in this study because the popularity of a YouTube product category as measured by distribution volume constitutes the theoretical criterion for case selection. The following variables, which are Twitter metadata, are considered the criteria of the study, because they act as proxies for the size of Twitter communities and the size of Twitter conversations: *Total_Tweets*, *Broadcast_Tweets*, *Engaged_Tweets*, *Community_Tweets*, *Total_People*, *Broadcast_People*, *Engaged_People* and *Largest_Community*. All are defined in Table 1.

III. ANALYSIS AND RESULTS

As shown in Table 2, the six chosen cases were categorized based on their popularity as measured by percentage of distribution volume of all videos on the YouTube platforms. They were binned into three categories: high, medium and low. Table 2 also displays cumulative numbers for metadata categories that were defined in Table 1.

According to our guiding proposition, products categorized as high in popularity were supposed to generate communities that were bigger in size, both in terms of number of tweets and people involved, than products that were categorized medium or small. However, Table 2 illustrates that this is not the case. According to all metadata metrics, which act as criteria for this study, the "Entertainment" category, which was categorized as 'high' based on YouTube popularity, generated fewer tweets than the "Comedy" and "Sports" categories, which were categorized as 'medium'. The numbers were not even close. For example, "Entertainment", generated 16,365 community tweets, whereas "Comedy" and "Sports" generated 25,624 and 32,778 community tweets, respectively, over the same period of time. This trend can also be extrapolated community sizes. For example, the largest communities within "Comedy" and "Sports" consisted of 24555 and 29998 participants, respectively, whereas "Entertainment" consisted of only 15,882. Large discrepancies in activity can also exist in communities that discuss product categories of comparable popularity. For example, the "Howto" category and the "Science" category exhibit comparable popularity, with respective distribution volume percentages of 3.1% and 2.86%. However, all metadata criteria indicate that the Twitter conversations about and Twitter communities associated with "Science" were more than four times larger than conversations about or communities associates with "Howto".

TABLE 2: METADATA OVERVIEW

Cases	Product Category	Popularity	% of All Videos	Total_ Tweets	Broadcast _Tweets	Engaged_ Tweets	Community _Tweets	Total_ People	Broadcast _People	Engaged_ People	Largest_ Community
1	Music	High	30.7	3,097,847	713,824	2,384,023	1,586,149	2,586,586	898,282	1,688,304	1,456,770
2	Entertainment	High	14.59	44,984	10,762	34,222	16,365	45,236	16,670	28,566	15,822
3	Comedy	Medium	5.2	94,111	33,350	60,761	25,624	83,175	37,456	45,719	24,555
4	Sports	Medium	6	129,182	67,476	61,706	32,778	77,617	25,776	51,841	29,998
5	Howto	Low	3.1	10,856	3,213	7,643	4,299	10,557	4,082	6,475	4,203
6	Science	Low	2.86	49,332	13,462	35,870	22,598	52,785	20,157	32,628	21,277

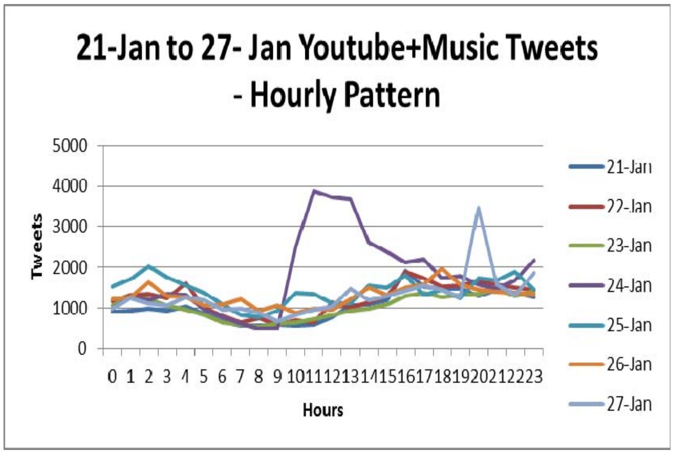

Figure 3: Hourly Patterns of Tweets between Jan. 21st, 2014 and Jan. 27th, 2014 in the Music Category

The collected tweets show a daily pattern of tweeting. For example, fig. 3 displays routinely occurring hourly patterns for data collected in the "Music" category between Jan. 21st, 2014 and Jan. 27th, 2014. However, fig. 3 also identifies some exceptions, which manifest themselves as "bumps" on Jan. 24th and Jan. 27th. These bumps are associated with the following events:

- 24th Seoul Music Awards -- Jan. 22nd, 2014
- 59th Filmfare Awards -- Jan. 24th, 2014
- 56th Annual Grammy Awards -- Jan. 26th, 2014

The impact (with delay) of the Seoul Music Awards, held on Jan. 22nd, can be seen on the tweet volume of Jan. 24th. The 24 hour pattern is consistent with previous large-scale studies undertaken on Twitter *(Dodds et al. [17])*. The 24-hour cycle started in accordance with Greenwich Meridian Time (GMT) in this study.

IV. Conclusions

As suggested by our guiding proposition, products categorized as high (in terms of popularity) on YouTube were supposed to generate Twitter communities that were bigger in size, than products that were categorized medium or small, both in terms of number of tweets and the number of people involved. The results of our study indicate that this does not hold true; the guiding proposition could not be confirmed. *A positive correlation between the popularity of a YouTube product category and the size of the Twitter conversation that the product category generates could not be established*, suggesting that activity does NOT necessarily transcend platforms. Evidently, the number of online conversations on a social networking platform about a specific product category on another media platform does not necessarily indicate how popular that product category is, and the size of the online community that discusses the product category is not necessarily an indicator of the product's popularity.

V. Discussion

A. Contributions to Management Practice

Due to the substantial impact of online social networks on marketing and e-commerce *(Weber & Mayande [52])*, these findings have significant implications for management practice. First and foremost, online social networks are disrupting traditional marketing models. Millions of consumers are continuously engaging in highly fluid conversations *(Dodds, et al. [16])*. As a consequence, the trends in society from which market needs for products and services are increasingly being articulated or even determined in cyberspace *(Chakrabarti & Berthon [14], Deighton [15])*. In addition, in activities such as marketing, customer service and product innovation, firms and organizations are increasingly able to take advantage of business ecosystems *(Afsarmanesh & Camarinha-Matos [1])* by leveraging their network value *(Bressler & Grantham [9])*. Furthermore, firms and organizations utilize online social networks to coordinate business and information exchanges because these networks are central to many successful business models *(Feller, et al. [21])*. *Weber & Mayande [52]* consequently argue that significant competitive advantage in the globalized economy of the 21st Century can be derived from understanding how online social networks are structured and how they behave.

B. Contribution to Theory

Unfortunately, practicing firms that are engaging with online social networks are unable to make sense of the phenomenon, in part because they are not able to rely on a solid theoretical foundation *(Aral, et al. [3], Li & Bernoff [34])*. Extant theory of social networks may not apply to online social networks because it is based on observations of the real world *(Mayande [37])*. Practicing firms also lack sufficient practical experience to comprehend how online social networks behave *(Wiertz, et al. [54])*. They may consequently really be grossly misallocating resources due to their nescience of the phenomenon *(Weber, Hasenauer & Mayande [51])*. Empirical confirmation of these assertions from the academic literature constitutes the primary theoretical contribution of this paper.

C. Limitations

Some of the limitations of this study may be a direct consequence of the research methods that have been deployed. For example, starting the 24-hour data collection cycle in accordance with Greenwich Meridian Time (GMT) is somewhat arbitrary. We consequently recommend further research to determine whether and how the results of this study are impacted by changing the start times of the 24-hour cycle. Furthermore, in this study, relevant Twitter communities were identified based on the presence of the word "YouTube" and the product category names in a tweet. A product category on YouTube, for example "Entertainment", might encompass various types of videos that do not fall under the conversations on Twitter in which the word "Entertainment" is used. For example, videos of movie trailers might be grouped under "Entertainment" category on YouTube but people talking about the movie trailers on Twitter might not use the word "Entertainment" in their tweet. This may partly be due to the limitations put forth by the platform itself (e.g., the 140 character limit on Twitter). However, this might not be the case for Music category. People engaged in conversations on Twitter about "Music" may use the word "Music" in all of their conversations. As a result, the "Music" conversation might generate one large cohesive community, while "Entertainment" may spawn multiple communities on Twitter *(Goggins and Petakovic [26])*. Therefore, further research is required to understand how community definitions translate across platforms.

D. Implications and Suggestions for Further Research

An in depth understanding of commercial activity in the 21st Century may hinge on investigating how online networks organize, behave and evolve. In particular, exploring how network structure and knowledge flows influence each other and how they impact a variety of network phenomena including influence may be of paramount importance *(Edwards [18] [19], Mayande [37], Weber & Mayande [52])*. Most importantly, these studies would have to identify the location of the most influential members of the network *(Aral & Walker [5])*. The results of such studies could help managers in real-world organizations design routines, structures, processes and practices from which radically innovative products and dramatically improved services can be developed *(Weber & Mayande [52])*. They should also be able to improve their ability to create social contagion through viral product design *(Aral & Walker [4])*.

The findings of this paper suggest that novel approaches to online marketing may be required to meet the needs of the modern consumer. First and foremost, today's marketers need better measures of influence in social networks. Centrality metrics from graph theory *(e.g., Freeman [22] [23])* can approximate the communication activity of a particular node (degree centrality); the control over the communication process is exerted by a node in the network (betweenness centrality); and the efficiency of a node's communication process (closeness centrality), but they do not really measure influence within the network *(Mayande [37], Weber & Mayande [52])*. Eigenvector centrality *(Bonacich [10])* measures the extent of a node's influence measures influence as a function of the spread of information very effectively *(Bonacich [11])*, but not as a function of how rapidly information spreads *(Mayande [37], Weber & Mayande [52])*. Entropy centrality *(e.g., Mayande & Weber [38], Nikolaev, et al. [43], Tutzauer [50], Weber, et al. [53])* appears to do that more effectively *(Weber & Mayande [52])*, in part because it measures the amount of information that can concentrate in a particular node *(Weber, et al. [51])*.

Recent research on online social networks also suggests that directionality may matter substantially *(Mayande [37], Mayande & Weber [39])*. A person that affects the propagation of information does not necessarily influence how it is consumed, and conversely *(Barley and Tolbert [7], Giddens [24], Goggins & Petakovic [26], Orlikowski [44])*. In order to identify key influencers within an online social network and deduce their behavioral traits, digital marketers consequently need to analyze patterns of consumption and propagation of information across the network *(Weber & Mayande [52])*.

Finally, as the study in this paper has shown, the behavior of online social networks may be more dependent on context than it is on scale. One cannot simply assume that more popular product categories generate larger conversations on social media, which implies, conversely, that analyzing the largest online conversations will not necessarily lead to the most popular product categories. Furthermore, recent research *(Mayande [37], Weber & Mayande [52])* has demonstrated that the structure of the network that surrounds the influencer impacts how information flows through it, and conversely. Both factors impact the loci of influence, and both can change radically over short periods of time *(Mayande [37])*. Successful online marketing will consequently consist of identifying the loci of highest influence in each network and exploring the context in which each network operates. Tools and methods for doing so are under development *(Weber, et al. [51])*, but further research must be conducted for these tools and methods to be effective *(Weber & Mayande [52])*.

REFERENCES

[1] H. Afsarmanesh and L. M. Camarinha-Matos, "A framework for management of virtual organization breeding environments," In *Collaborative Networks and their Breeding Environments, (PRO-VE'05)*, Valencia, Spain: IFIP — The International Federation for Information Processing, Vol. 186, Springer, 2005, pp. 35-48.

[2] http://preview.alexa.com/siteinfo/youtube.com (accessed 04/09/2014).

[3] S. Aral, C. Dellarocas and D. Godes, "Introduction to the special issue-social media and business transformation: A framework for research," *Information Systems Research,* Vol. 24, 2013, pp. 3-13.

[4] S. Aral and D. Walker, "Creating social contagion through viral product design: A randomized trial of peer influence in networks," *Management Science,* Vol. 57, 2011, pp. 1623-1639.

[5] S. Aral and D. Walker, "Identifying influential and susceptible members of social networks," *Science,* Vol. 337, 2012, pp. 337-341.

[6] C. Ayres, "Revenge is best served cold – on YouTube: How a broken guitar became a hit," *The Times* (UK), July 22, 2009.

[7] S. R. Barley and P. S. Tolbert, "Institutionalization and structuration: Studying the links between action and institution," *Organization Studies,* Vol. 18, 1997, pp. 93-117.

[8] H.-P. Blossfeld and G. Rohwer, *Techniques of event history modeling: New approaches to causal analysis (2nd ed.)*, Mahwah, NJ, USA: Lawrence Erlbaum Associates Publishers, 1995.

[9] S. E. Bressler and C. Grantham, *Communities of Commerce: Building Internet Business Communities to Accelerate Growth, Minimize Risk, and Increase Customer Loyalty,* New York, NY, USA, McGraw-Hill, Inc., 2000.

[10] P. Bonacich, "Factoring and weighing approaches to status scores and clique identification," *Journal of Mathematical Sociology,* Vol. 2, 1972, pp. 113-120.

[11] P. Bonacich, "Some unique properties of eigenvector centrality," *Social Networks,* Vol. 29, 2007, pp. 555-564.

[12] M. Cha, H. Haddadi, F. Benevenuto and P. K. Gummadi, "Measuring user influence in Twitter: The million follower fallacy," *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, Vol. 10, 2010, pp. 10-17.

[13] M. Cha, A. Mislove and K. P. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network,"

Proceedings of the 18th international conference on World Wide Web, 2009. ACM, pp. 721-730.

[14] R. Chakrabarti and P. Berthon, "Gift giving and social emotions: Experience as content," *Journal of Public Affairs,* Vol. 12, 2012, pp. 154-161.

[15] J. Deighton, "What are marketers' top concerns?" (MSI Research Priorities 2012-14), Jun 25, 2012 ed.: Marketing Science Institute.

[16] P. S. Dodds, R. Muhamad and D. J. Watts, "An experimental study of search in global social Networks," *Science,* Vol. 301, 2003, pp. 827-829.

[17] P. S. Dodds, K. D. Harris, I. M. Kloumann, C. A. Bliss and C. M. Danforth, "Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter," *PloS one,* Vol. 6, Dec. 7, 2011, e26752.

[18] J. Edwards, "DATA: Google totally blows away Facebook on ad performance," *Business Insider,* May 15, 2012, 6:00 AM ed.

[19] J. Edwards, "Priceline CEO, who has a $1.8 billion online ad budget, says Facebook and Twitter are useless," *Business Insider*, April 15, 2014, 9:42 AM.

[20] K. M. Eisenhardt, "Building theories from case-study research," *Academy of Management Review*, Vol. 14, pp. 532-550, 1989.

[21] J. Feller, P. Finnegan, B. Fitzgerald and J. Hayes, "From peer production to productization: A study of socially enabled business exchanges in open source service networks," *Information Systems Research,* Vol. 19, 2008, pp. 475-493.

[22] L. Freeman, "A set of measures of centrality based on betweenness," *Sociometry,* Vol. 40, 1977, pp. 35-41.

[23] L. C. Freeman, "Centrality in social networks: Conceptual clarification," *Social Networks,* Vol. 1, 1979, pp. 215-239.

[24] A. Giddens, *The Constitution of Society: Outline of the Theory of Structuration*, Berkeley and Los Angeles, CA, USA: University of California Press, 1986.

[25] http://gigaom.com/2012/06/04/twitter-shows-when-we-tweet-and-explains-why-its-search-sucks/ (accessed on 4/27/2014).

[26] S. P. Goggins and E. Petakovic, "Connecting theory to social technology platforms: A framework for measuring influence in context," *American Behavioral Scientist*, Vol. 58, No. 10, 2014, pp. 1376-1392.

[27] B. Gonçalves and J. J. Ramasco, "Human dynamics revealed through Web analytics," *Physical Review E,* vol. 78, 2008, 026123.

[28] A. Halavais, "Bigger sociological imaginations: Framing big social data theory and methods," *Information, Communication & Society*, Vol. 18, No. 5, 2015, pp. 583-594.

[29] C. Honeycutt and S. Herring, Beyond Microblogging: Conversation and Collaboration via Twitter. *Hawaii International Conference on System Sciences,* 0, 2009, pp. 1-10.

[30] B. A. Huberman, D. M. Romero, and F. Wu, "Social Networks that Matter: Twitter Under the Microscope," *CoRR,* 2008, abs/0812.1045.

[31] B. Jansen, M. Zhang, K. Sobel and A. Chowdury, "Twitter power: Tweets as electronic word of mouth," *Journal of the American Society for Information Science and Technology,* Vol. 60, 2009, pp. 2169-2188.

[32] D. Leonard-Barton, "A dual methodology for case studies: Synergistic use of a longitudinal single site with replicated multiple sites," *Organization Science,* Vol. 1, 1990, pp. 248-266.

[33] M. Khammash and G. H. Griffiths, "'Arrivederci CIAO.com, Buongiorno Bing.com'—Electronic word-of-mouth (eWOM), antecedences and consequences," *International Journal of Information Management,* Vol. 31, 2011, pp. 82-87.

[34] C. Li and J. Bernoff, *Groundswell: Winning in a World Transformed by Social Technologies,* Boston, MA, USA: Harvard Business Press, 2008.

[35] A. Lindsay, J. Kaykas-Wolff and C. Mathwick, "Key influencers: Locating, measuring and creating programs to influence social media influencers, *Proceedings of the Academy of Marketing Science (AMS) Annual Conference*, 2010, p 1.

[36] P. Longart, What drives word-of-mouth in restaurants? *International Journal of Contemporary Hospitality Management,* Vol. 22, 2010, pp. 121-128.

[37] N. Mayande, *Network Structure, Network Flows and the Phenomenon of Influence in Online Social Networks: An Exploratory Empirical Study of Twitter Conversations about YouTube Product Categories*, Dissertation in Engineering and Technology Management, Portland State University, July 2015. http://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=3471&context=open_access_etds

[38] N. Mayande and C. M. Weber, "Designing virtual social networks for for-profit open innovation," *Proceedings of PICMET '11*, Portland, Oregon, July 31-Aug. 4, 2011.

[39] N. Mayande and C. M. Weber, "Directed interaction networks and their impact on social media," *Proceedings of PICMET '13*, San Jose, CA, USA, July 28 to August 1, 2013, pp. 1709-1722.

[40] M. R. Meiss, F. Menczer, S. Fortunato, A. Flamini and A. Vespigiani, "Ranking web sites with real user traffic," *Proceedings of the ACM 2008 International Conference on Web Search and Data Mining,* Palo Alto, California, USA, Sept. 7-11, 2008, pp. 65-76

[41] S. Milstein, A. Chowdhury, G. Hochmuth, B. Lorica and R. Magoulas, 2008. *Twitter and the micro-messaging revolution: Communication, connections, and immediacy—140 characters at a time,* Sebastopol, CA, O'Reilly Media, 2008.

[42] S. Ng and S. Vranka, "P&G Shifts Marketing Dollars to Online, Mobile," *Wall Street Journal*, Aug. 1, 2013, 8 PM EDT.

[43] A. G. Nikolaev, R. Razib and A. Kucheriya, "On efficient use of entropy centrality for social network analysis and community detection," *Social Networks*, Vol. 40, Jan. 2015, pp. 154–162.

[44] W. J. Orlikowski, "Using technology and constituting structures: A practice lens for studying technology in organizations," *Organization Science,* Vol. 11, No. 4, 2000, pp. 404-428.

[45] D. Ramage, S. Dumais and D. Liebling, "Characterizing Microblogs with Topic Models," *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, Washington, DC, USA, May 23-26, 2010.

[46] http://www.statisticbrain.com/twitter-statistics/ (accessed 04/09/2014)

[47] http://www.sysomos.com/reports/youtube/#categories (accessed on 04/09/2014)

[48] J. Teevan, D. Ramage and M. R. Morris, "#TwitterSearch: A comparison of microblog search and web search," *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*, Hong Kong, China, Feb. 9-12, 2011, pp. 35-44.

[49] M. Thelwall, P. Sud and F. Vis, "Commenting on YouTube videos: From Guatemalan rock to El Big Bang," *Journal of the American Society for Information Science and Technology*, Vol. 63, 2012, pp. 616-629.

[50] F. Tutzauer, "Entropy as a measure of centrality in networks characterized by path-transfer flow," *Social Networks*, Vol. 29, No. 2, May 2007, pp. 249–265.

[51] C. M. Weber, R. Hasenauer and N. V. Mayande, "Quantifying nescience: A decision-aid for practicing managers," *International Journal of Innovation and Technology Management,* forthcoming, 2018.

[52] C. M. Weber and N. V. Mayande, "Knowledge Flows and Influence in Online Social Networks: Proposing a Research Agenda," *Proceedings of PICMET '17*, Portland, Oregon, July 9-13, 2017, pp. TBD.

[53] C. Weber, V. Sankaran, K. Tobin and G. Scher, "Quantifying the value of ownership of yield analysis technologies," *IEEE Transactions of Semiconductor Manufacturing,* Vol. 15, No. 4, Nov. 2002, pp. 411-419.

[54] C. Wiertz, C. Mathwick, K. De Ruyter and B. Dellaert, "A balancing act: Exploring governance in a virtual P3 community," *Advances in Consumer Research,* Vol. 37, 2010, pp. 672 - 673.

[55] S. A. Williams, M. M. Terras and C. Warwick, "What do people study when they study Twitter? Classifying Twitter-related academic papers," *Journal of Documentation,* Vol. 69, 2013, pp. 384-410.

[56] C. Xu, , C. Dale and L. Jiangchuan, "Statistics and Social Network of YouTube Videos," Quality of Service, 16th International Workshop on, June 2-4, 2008, pp. 229-238.

[57] R. K. Yin, *Case Study Research: Design and Methods* (2nd ed.), Beverly Hills, CA, Sage Publishing, 1984.

[58] D. Zhao and M. Rosson, "How and why people Twitter: The role that micro-blogging plays in informal communication at work," *Proceedings of the ACM 2009 International Conference on Supporting Group Work*, Sanibel Island, Florida, USA, May 10-13, 2009, pp. 243-252.