



Journal of Information Technology Management

ISSN #1042-1319

A Publication of the Association of Management

EXPLORING SOCIAL MEDIA ANALYTICS ON COMMUNITY DEVELOPMENT PRACTICES

DI SHANG

LONG ISLAND UNIVERSITY BROOKLYN

Di.ShangGhriga@liu.edu

MOHAMMED GHRIGA

LONG ISLAND UNIVERSITY BROOKLYN

Mohammed.Ghriga@liu.edu

ABSTRACT

The use of social media and social networks has become a significant community development tool to reach a wide audience, share information in real-time, improve communications, and build relationships. In this research, we aim to demonstrate how social media analytics provide a valuable set of techniques for assessing community development and identifying strategies to develop best practices. We analyze the activities and structure of a local community - Brooklyn Tech Triangle based on data collected from Twitter. Our results suggest that community development practitioners can use social media to promote the fundamental tenets and promote their initiatives in the community. Furthermore, practitioners can apply social media analytics to discover strategies for keeping leadership roles, connecting the community and identifying value-added influencers. Our findings show the potentials social media analytics offer for practitioners in improving their community development practices.

Keywords: social media, social networks, social media analytics, community development, network analysis

INTRODUCTION

Social media is transforming how companies communicate with their customers and stakeholders, and it forms a major part of most company's communication and marketing strategies. Previous research [18] concludes that social media differs from traditional media in three ways: 1) social media facilitates two-way interaction with a large number of people at the same time; 2) companies and customers are connected with each other directly; and 3) social media platforms are free to join and cost efficient for marketing and communication. Companies can glean immediately what their target audience wants and who the

key influencers are, and thus they can adjust their strategies to meet their customer's specific needs.

The use of social media and social networks has also become a significant community development tool to reach a wide audience, share information in real-time, improve communications, and build relationships. Social media data provide the opportunity for understanding how a community is formed and maintained online. With information collected from social media, practitioners can actively engage in community development and address community issues. Moreover, on social media platforms, information can be broadly disseminated, and greater transparency and accountability can lead to increases of trust within community members [4]. Through network-

ing, social media also enhances the capacity of community leaders and key influencers within the community.

Meanwhile social media also raised challenges to community development that resulted from its great interconnectedness. Community development practitioners are working in more dynamic and boundless network-based settings, and they have to rethink the traditional practices applied to stable physical location-based communities. In addition, social media changed the ways people communicate with others and access and share information. People freely exchange information and express opinions, which undercuts centralized control of attitudes and behaviors. Therefore, researchers [e.g., 13] suggest that community development practitioners need to recognize and scrutinize how social media can challenge established ways of practicing community development.

In this research we aim to demonstrate how social media analytics can provide a valuable set of techniques for analyzing community development and discovering best practices from social media data. We examine the activities and structure of the social media presence of a local community - Brooklyn Tech Triangle based on data collected from Twitter on the key practitioners. We aim to answer the following specific questions:

- How social media analytics help to assess community development practitioners' use of social media to promote the fundamental tenets of the community?
- How social media analytics help to discover practices for enhancing community development?

We visualize the engagements of community members and network properties of the community. Results of our analysis show that community development practitioners can promote an active community centered in their initiatives and keep leadership roles in connecting the local community through social networks. In addition, community leaders can apply social media analytics to discover strategies to keep leadership roles by identifying and utilizing value-added influencers to transfer information through social networks.

The rest of the paper is structured as follows: section 2 discusses the related work done in this area. Sections 3 describe the context and methodology. In section 4, we present our findings, and we provide discussions and recommendations based on our results in section 5. Section 6 concludes the paper.

RELATED WORKS

For the last few years, interest in social media and social networks has increased and become an active

research area due to commercial and social importance, the potential of applications, and the value of analysis. As a result, organizations understand the importance of utilizing social media to exist in today's technologically advanced market environment. New algorithms, models, and applications have been developed to address particular domains and resolve distinct problems in social media [14].

As social media becomes well recognized for its potential as a social communication channel, many companies have used tools such as Facebook, Twitter, and blogs for communication and engagement with their customers as well as employees [9]. These social platforms act as a connective space for communication and information sharing. Customers become active contributors to content creation contrary to being passive audience in traditional marketing [3]. Social media usage has contributed to communication practices in both co-located and geographically dispersed networks, and by engaging the audience via interactions and exchanges social media empowers organizations to establish constant positive relationships with customers [19].

Social media enables companies to network with customers in order to build relationships and achieve a better understanding of customer needs. For instance, analyzing twitter data helps institutions or companies to have better understanding about their customers' feedbacks and opinions, attitudes, perceptions and behavior [3]. Organizations and policy makers can benefit from targeting members to effectively spread valuable information [4]. Organizations nowadays focus on analyzing social networks to know who is central in these networks and identifying members who are best placed to diffuse information [2].

The use of social media and social networks has also become a significant community development tool to reach a wide audience, share information in real-time, improve communications, and build relationships [7]. Social media and social networks provide diverse means of communication and engagement to support community development efforts [16]. Community participants can use social media to generate content and opinions, create an online identity and build networks, and they can actively engage in learning and understanding about community issues with information on social media platforms. In addition, through networking community practitioners can enhance their leadership capacity within the community [10]. Therefore, it is important that community leaders closely follow best practices in order to have effective communication through social media platforms [1].

Social media analytics "is concerned with developing and evaluating informatics tools and frameworks to

collect, monitor, analyze, summarize, and visualize social media data ... to facilitate conversations and interactions ... to extract useful patterns and intelligence...” [20]. Social media data provide the opportunity for understanding how a community is formed and maintained online. For instance, visualization of social networks can illustrate the interconnections of community members in the networks. In addition, network graphs can also show how the community is centralized on certain subgroups or otherwise isolated from each other. In addition, social media sentiment analysis can be an excellent source of information and can provide insights that improve a practitioner’s success of community development and community supports. Made use of these sentiments open up new possibilities and opportunities. Thus, integrating social media analytics into community development practices is a useful way to assess the impacts of community development works [7].

CONTEXT AND METHODOLOGY

This research focuses on a local community - Brooklyn Tech Triangle in understanding the use of social media analytics for community development practices. Brooklyn is the largest of New York City’s five boroughs by population and the second largest by area. The Brooklyn Tech Triangle is made up of the following three key partners: Downtown Brooklyn, DUMBO, and the Brooklyn Navy Yard. In the past few years, it has emerged as one of the largest tech sectors in New York City, attracting entrepreneurs and startup companies to the area since most innovation companies want to be close to each other to share ideas and benefit from collective expertise. The Brooklyn Tech Triangle initiative is an effort to bring together support for growing the innovation economy of the borough. The Downtown Brooklyn is missioned to attract new investment and improve the business environment for existing companies by developing infrastructure and promoting a cohesive community. DUMBO, a neighborhood located between the Brooklyn and Manhattan Bridges, is Brooklyn’s epicenter of art and creativity. The DUMBO Improvement District is founded to enhance the quality of life in the area, and it initiates neighborhood planning and advocacy to amplify the area’s creative and innovative vibe. The Brooklyn Navy Yard Development Corporation manages the Brooklyn Navy Yard on behalf of the City of New York. It aims to create quality jobs in the modern industrial sector and connect local community to resources and economic opportunities. The Yard has become a national model for sustainable urban industrial parks through development of infrastructure and green

buildings which host tech-driven manufacturers and companies¹.

Since the use of social media and social networking has become an important approach in social interaction and economic development, the three practitioners all have their individual social media presences to promote an active and cohesive community, attract entrepreneurs and investment, and connect the local community with the economic opportunity and resources. Twitter enables prompt but relatively shallow conversations, and its simplicity is important for fast news sharing and just-in-time updates [12]. Using related keywords (Twitter hashtags as shown in Table 1), a total 4528 tweets were collected from Twitter through dashDB on IBM’s BlueMix platform (a data warehousing solution on the cloud) from April 2014 to November 2016, over a thirty-two-month span.

Table 1: Twitter Data of Brooklyn Tech Triangle

Search Term	Number of Tweets
#DowntownBrooklyn	1832
#DumboBid	1308
#BrooklynNavyYard	1388
Total	4528

To understand how social media facilitates the formation of relationships among community members, Ang [1] proposes four pillars to achieve it: (1) connectivity, (2) conversations, (3) content creation and (4) collaboration. Statistics about community participation can provide important insights about the conversation and content creation of a community. Meanwhile, network analysis can help describing the connectivity and collaboration of participating members, identifying the patterns of interactions, and tracing how information flows within the community. These can be particular key measures of the health of a community in terms of reducing isolation and disconnectedness and receiving quick responses from community members. Therefore, community engagement, topic analysis, sentiment analysis and network analysis are all important metrics that community practitioners should seek to understand.

We analyze the Twitter data set using R packages, including *RBGL*, *SNA*, *TM*, and *Wordcloud*. R is an open source programming language and environment for

¹ <http://brooklyntechtriangle.com/about/> Accessed December 6, 2018

statistical computing, and it is also a great tool for data visualizations [21]. For performing visualizations of the data analysis, we utilize following packages in R, *ggplot2*, *Plotly* and *RColorBrewer*.

ANALYSIS AND RESULTS

We first examine the content creation of community users and community development practitioners in using Twitter to promote the fundamental tenets of the community. As shown in Figure 1, the community development practitioners are active in term of content and

traffic generation in the community. The number of tweets posted or retweeted by practitioners accounts for more than 10% on average of the total number of tweets generated by the community over time. In addition, large portion (around 56% on average) of these tweets are original posts, which indicates active content creation of community users. Whilst, 44% of the tweets posted was retweeted or shared by users, which indicates that the community also have large following users who share content and drive traffic.

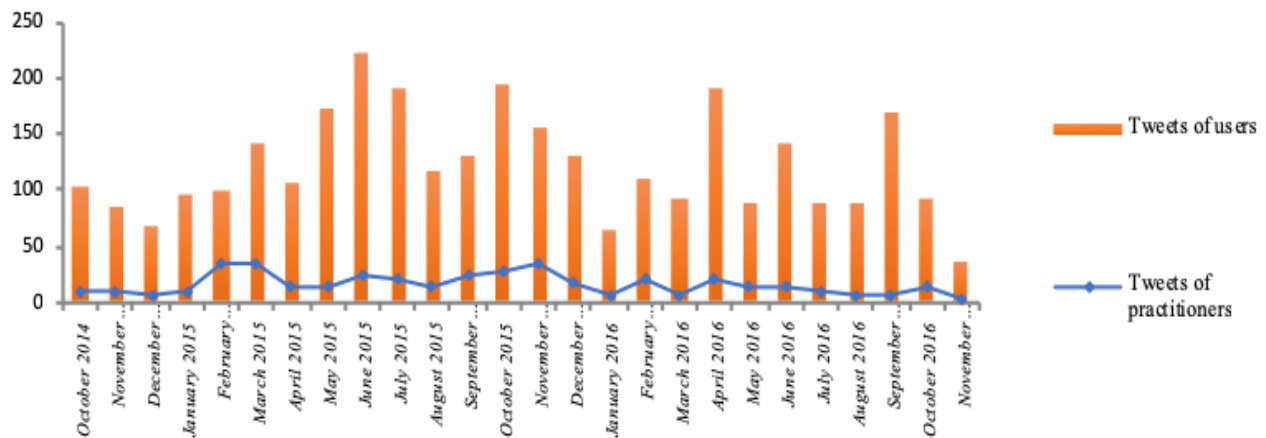


Figure 1: Number of Tweets Contributed by Practitioners and Users

Understanding the patterns of community users' active behavior help practitioners to develop practices for promoting tenets and sharing information [5]. Figure 2 shows the volume of tweets posted and shared by active users of the community in the day of a week over the three-year period (as shown in the left chart). Our analyses reveal that through the week the engagement of users in the activity of posting and sharing was peaking on Thursday, and next to it we observed Wednesday and Friday. It suggests that the best period of the week the community users are active in posting and sharing the tweets is towards the end of the week. Examining the engagement of active users of the community by hour, we observe the

raise of activity (posting/sharing) from 9.00 AM to 3.00 PM, and from 3.00 PM there are fluctuations in the activity and this continues with the raise of few peaks, till 9.00 PM (as shown in the right chart). This suggests that the users are likely to tweet during lunch breaks in the afternoon where highest peak of activity is observed at 3.00 PM, and at the time of commute in the evening from 5.00 PM and then at night before 9.00 PM. During nighttime from 9.00 PM we observe very less activity till the next morning. The above findings can help community practitioners identify the best time slots to communication and share information with community members.

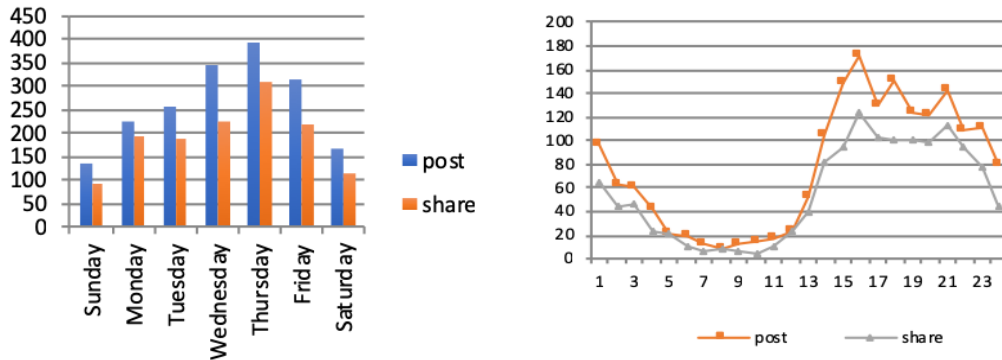


Figure 2: Volume of Tweets Posted and Shared by the Day of a Week (left) and by Hour (right)

One important community development practice of the practitioners is using social media to promote events that attract entrepreneurs and investment and connect the local community with economic opportunity and resources. Based on word cloud and human annotation (by graduate students) [6], we develop a domain-specific dictionary of events in the community over the period of time. Our topic analysis shows that around 40% of the tweets are related to community events, which is consistent to the community practitioners' initiative to connect with Brooklyn's best and brightest by hosting events

and building entrepreneurs networks. Figure 3 shows the average number of retweets within 24 hours of the announcement of events in the community. The announcements attracted around 2980 retweets on average within 24 hours, and it suggests that community users quickly respond to community events and disseminate the information in the community. The number of retweets sharply decreased after 24 hours, which also reflects the fact that trending topics on Twitter can quickly change from one term to another [15].

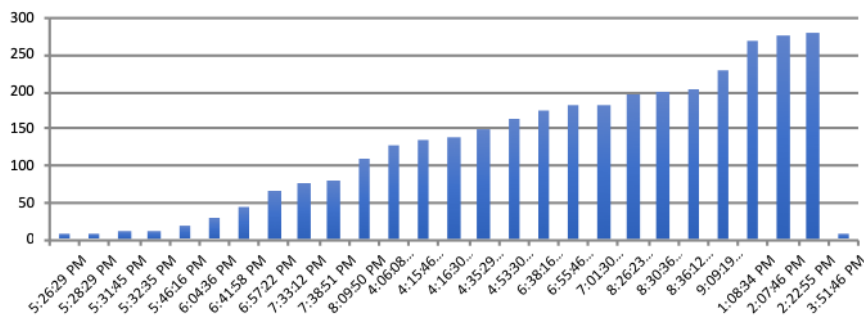


Figure 3: Average Number of Retweets within 24 hours of the Announcement of Events

Sentiment analysis is the process that “aims to determine the attitude of a speaker or a writer with respect to some topic” [11]. Sentiment analysis empowers organizations by providing extensive, insightful information regarding their target audiences' sentiments. The benefit of community practitioners tracking the sentiment of community users is that it allows quantifying perceptions about their leaderships, community initiatives, engage-

ment and campaigns. Community practitioners can use the information for devising effective communication strategies and can therefore, improve their reputation and leadership. We used *text2vec* and *glmnet* R packages to train a model with 1.6 million labelled tweets, and we conducted sentiment scoring of our Twitter data using the trained model. The model generated sentiment scores of every tweet in a range of values from 0 (completely negative) to

1 (completely positive). We cut off the scores as: negative from 0 to 0.35, positive from 0.66 to 1, and neutral 0.36 to 0.65. Human annotation of 500 sampled tweets suggests that the model achieved an accuracy of 91%. Our

results (as shown in Figure 4) suggest that community users generally express positive sentiment during the period of time.

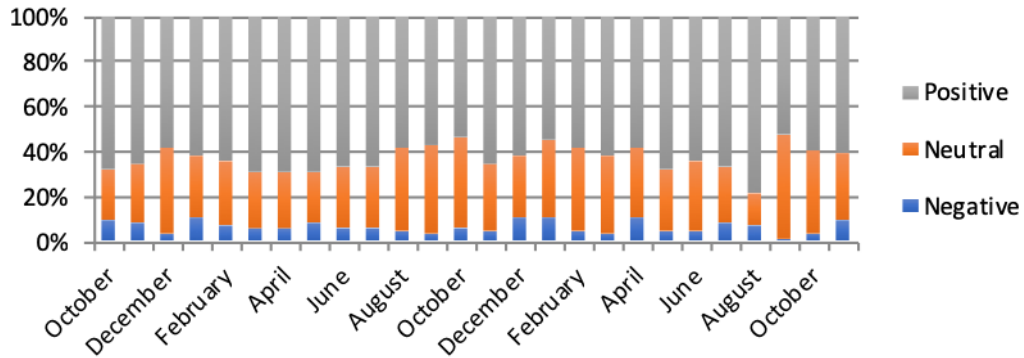


Figure 4: Sentiment of tweets over time

Communities can be viewed as sets of individuals and organizations and their relationships, and community development is strengthening and extending networks of relationships between individuals, and between organizations. We next examine whether the community development practitioners present leaderships in the community and enhance the leadership capacity through networking.

Hubs are individuals in a network that are highly sought-after by other network members. Hubs of influence in a network are measured using indegree centrality - counts how many relationships point towards an individual [8]. Bridgers are individuals in a network who have connections to different clusters. Bridgers in a network provide valuable opportunities for growth and impact because they have access to perspectives and networks that are otherwise unknown to most network members. Finding bridgers in a network is typically done with the calculation called betweenness centrality [8]. Identifying members with high betweenness centrality score is important because it could help spread information in the social network faster, and it could also help protect the network from breaking [2]. A node with high betweenness centrality has a large influence on the transfer of information through the network, under the assumption that information transfer follows the shortest paths. Recognizing where and how information flows through networks is useful for determining how to strategically access it. As shown in Table 2, all of Downtown Brooklyn, DUMBO and Navy Yard ranked top in the list reflecting their leadership and influence on the social network.

Table 2: Centrality Scores of Key Accounts

Account	Indegree centrality	Betweenness centrality
DowntownBklyn	403	190966.7
DUMBOBID	229	118769.4
Navy Yard	44	40395.7
TechTriangleU	29	29458.0
Brownstoner	20	14137.4
Explore Brooklyn	51	11372.7
SFC Entrepreneur	35	11864.3
makeitinbk	18	10766.3

Network graph of the whole data set presents a large graph with many small clusters. Using *igraph* R package, we exclude these small clusters in order to reduce dimensionality. The most influential accounts in the network are labelled in the pruned network graph (Figure 5). In addition to the practitioners, certain Twitter accounts in the network play important roles in keeping the community active and the transfer of information through the network. For example, Brownstoner, a real estate broker, is active in tweeting and blogging about the development of Brooklyn, and such users play unexpected roles in connecting the community and transferring information through social networks. TuckerDBP, an advocator, is active in tweeting and blogging about the development of Brooklyn; and SFC, a college’s entrepreneurship center,

promotes entrepreneur and practitioner engagement with resources through its Open Virtual Incubator. Those value-added influencers promote the community develop-

ment initiative and enhance the leadership of the community development practitioners.

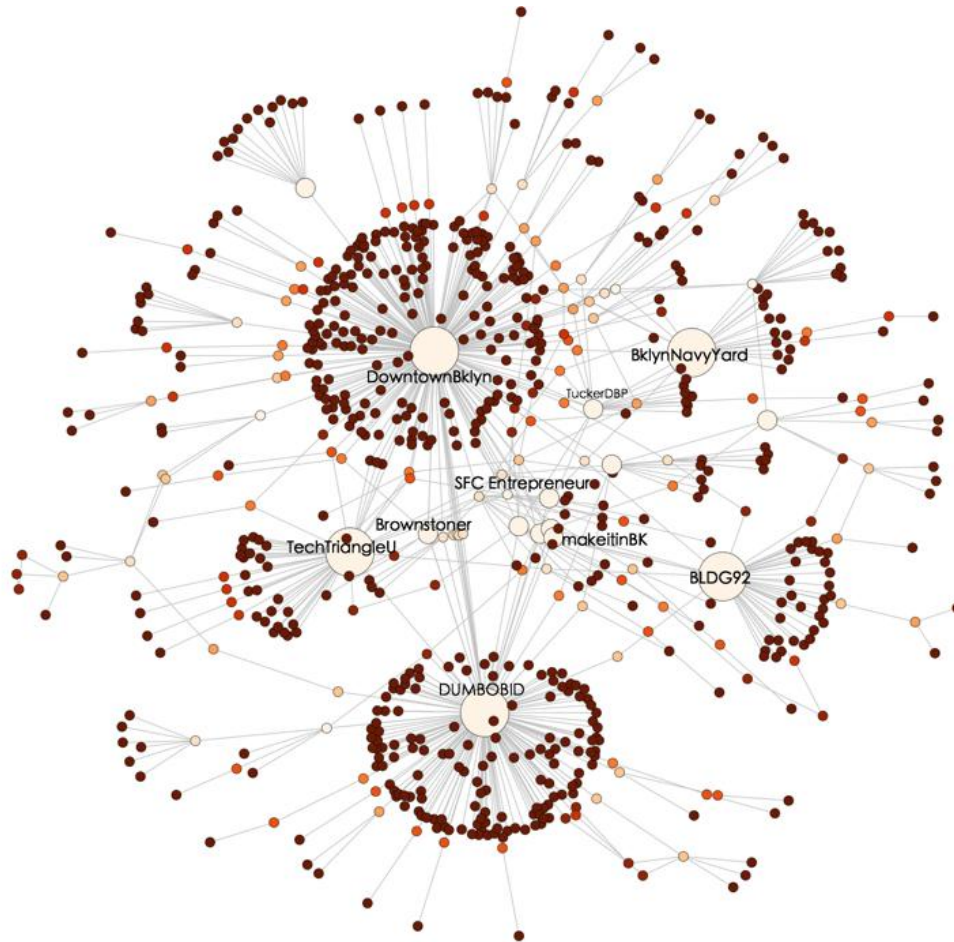


Figure 5: Pruned Network Graph of the Twitter Data Set

Network analysis can also provide information about how the network expand and grow more interconnected over time. We observed that the networks grow in terms of both size and density from 2014 to 2016. The network graph helps us to identify emergence of influencers over time. For instance, ‘makeitinbk’- Make It in Brooklyn, emerged as one of the important hubs in the network in 2015. ‘makeitinbk’ is an initiative to connect with Brooklyn’s best and brightest by hosting Innovation Summit and build entrepreneurs network by hosting pitch contests. It celebrates the entrepreneurial ethos, the grit, and the drive of the people that make Brooklyn. In 2016, “BE.IN” (The Brooklyn Education Innovation Network)

became a major hub and bridger in the network. BE.IN is a consortium of 10 institutions, 66,000 students, 11,000 faculty and staff that takes a comprehensive and innovative approach to collaborations of colleges and universities. Such information is helpful for practitioners to closely follow the change of network structure and thus employ appropriate influencers in connecting community.

DISCUSSION

Our analysis highlighted how social media is recognized as an important tool in community development. The range of tools available to practitioners are

varied and can provide significant insights into tracking community development, helping practitioners to make informed business decisions, and promoting the essential tenets of the community. Even though Brooklyn Tech Triangle and its key partners cannot control what people talk about on social media, knowing what they are talking about gives a good indication of the effectiveness of their development practices. Community practitioners can customize their online messages, social media strategies and campaigns in order to garner more positive and interesting topics. In our study the Brooklyn community practitioners serve the community through a collective leadership. As a successful collective leadership network relies on long-term preservation of the group's mission, online communities must manage participation for long-term effectiveness [10]. Community practitioners can apply social media analytics to monitor the patterns of community users' content contribution and active behavior to develop best practices for promoting tenets and sharing information, as demonstrated in our study.

Social media analytics provide the methods to investigate the relationships between community members and practitioners. Our analysis showed that the practitioners stay connected to the community and those key influencers have influence on the transfer of information through the social network. While institutional arrangements and participation of practitioners are important, we should also examine how users are networked. It is important to recognize that key individuals in community and their 'brokering' work in empowering local communities [17]. In our study influencers, such as Brownstoner, TuckerDBP and SFC, played expected or unexpected 'brokering' roles in connecting the community and transferring information through social networks. Our findings suggest that community development practitioners should identify active influencers in the network and utilize them to enhance their leadership and information sharing and communication, along with strategic alliances and collaborations.

Social media data provide valuable information for organizations to conduct text mining, social network analysis, opinion detections, and trends analysis. Techniques used in our research and some of our findings can be useful for practitioners to design platforms that use social media as a means of community development tracking. Such a system could perform tracking task with minimal supervision and with the ability to provide time-critical alerts and decision-support to community development practitioners. Social media data can be collated from different data sources and stored in a data repository. Data relevant to a particular interest can be retrieved, and different analysis techniques (namely, topic analysis, sen-

timent analysis, and network analysis) would operate on the retrieved data. The output of these analyses will be combined together in a trend analysis to track emerging or prevalent trends over time associated with community development. Community practitioners can also use the system to track new topics, assess the community user's attitude and identify the key influencers and the role they can play in the communication. An insight generation and reporting system will consolidate different forms of analytic results together to produce reports pertaining to insights for identifying key influencers, tracking early signs of new topics, or assessing community users' sentiments and concerns over issues.

CONCLUSION

Social media analytics provide a framework to research community development practices and help community practitioners to understand the effects of those practices on the relationships between organizations and members that constitutes a community. Using the case of Brooklyn Tech Triangle, we demonstrate how social media analytics provides a valuable set of techniques for analyzing community development information from social media data. While social media is considered a less formal platform for rapid release of critical information, our analysis provide evidence that community development practitioners can potentially benefit from utilizing social media as new tools to improve their communication strategies. In addition, our approach and the techniques used can apply to the analysis of other communities and help practitioners to design platforms using social media to track community development. The visual and metric outputs of social media analysis can be used by community practitioners to enhance the accomplishment of their organizational missions and strengthen their leadership. Social media analytics shows capabilities as effective methods to understand the efficacy of community development.

Our findings are based on analysis of a local community that has a relatively small Twitter data set. We can overcome the limitation in our future work by investigating more communities and also include data from other social media sources. More research on social media analytics applied in different contexts are needed to enable us to evaluate our metrics and validate the analysis and interpretation of the data. It is worth mentioning that tracking community development using social media also presents many challenges, such as real-time analysis of large volume of data, high level of background noise, and dynamic network structure. However, the underlying value coupled with challenges of using social media warrants future re-

search. In conclusion, our research exhibits that social media becomes a common venue for community developments. Community development practitioners can successfully promote community initiatives and influence community members by engaging in social media platforms. Social media analysis generates insightful information about community development, and evaluating the online activity can give community leaders' great advantage in not only building its online presence but understanding its target audience and solve issues within the community.

REFERENCES

- [1] Ang, L. "Community Relationship Management and Social Media," *Journal of Database Marketing and Customer Strategy Management*, Volume 18, Number 1, 2011, pp.31–38.
- [2] Bakshy, E., Hofman, J. M., Mason, W. A., and Watts, D. J. "Everyone's an Influencer: Quantifying Influence on Twitter," *Proceedings of the fourth ACM international conference on Web search and data mining*, 2011, pp.65–74. Hong Kong, China: ACM Press.
- [3] Burton, S., and Soboleva, A. "Interactive or Reactive? Marketing with Twitter," *Journal of Consumer Marketing*, Volume 28, Number 7, 2011, pp.491–499.
- [4] Chau, M., and Xu, J. "Business Intelligence in Blogs: Understanding Consumer Interactions and Communities," *MIS Quarterly: Management Information Systems*, Volume 36, Number 4, 2012, pp.1189–1216.
- [5] Chen, A., Lu, Y., Chau, P. Y. K., and Gupta, S. "Classifying, Measuring, and Predicting Users' Overall Active Behavior on Social Networking Sites," *Journal of Management Information Systems*, 2014, Volume 31, Number 3, 2014, pp. 213–253.
- [6] Deng, S., Sinha, A. P., Zhao, H., and Wu, J. "Learning Domain-specific Lexicon for Sentiment Analysis of Social Media Texts," *Proceedings of 23rd Workshop on Information Technology and Systems: Leveraging Big Data Analytics for Societal Benefits*, WITS 2013.
- [7] Ennis, G., and West, D. "Exploring the Potential of Social Network Analysis in Asset-based Community Development Practice and Research," *Australian Social Work*, Volume 63, Number 4, 2010, pp.404–417.
- [8] Freeman, L. C., Roeder, D., and Mulholland, R. R. "Centrality in Social Networks: ii. Experimental Results," *Social Networks*, Volume 2, Number 2, 1979, pp.119–141.
- [9] Gupta, V., and Gupta, M. "Social Media Mining: A New Framework and Literature Review," *International Journal of Business Analytics*, Volume 3, Number 1, 2016, pp.58–68.
- [10] Hoppe, B., and Reinelt, C. "Social Network Analysis and the Evaluation of Leadership Networks," *Leadership Quarterly*, Volume 21, Number 4, 2010, pp.600–619.
- [11] Hu, M., and Liu, B. "Mining and Summarizing Customer Reviews," *Proceedings of the 2004 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2004, p.168-177.
- [12] Jansen, B. J., Zhang, M., Sobel, K., and Chowdury, A. "Twitter Power: Tweets as Electronic Word of Mouth," *Journal of the American Society for Information Science and Technology*, Volume 60, Number 11, 2009, pp.2169–2188.
- [13] Kenny, S. "Community Development Today: Engaging Challenges through Cosmopolitanism?" *Community Development Journal*, Volume 51, Number 1, 2016, pp.23–41.
- [14] Leonardi, P. M. "Social Media, Knowledge Sharing, and Innovation: Toward a Theory of Communication Visibility," *Information Systems Research*, Volume 25, Number 4, 2014, pp.796–816.
- [15] Lin, J., and Mishne, G. "A Study of "Churn" in Tweets and Real-Time Search Queries," *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, 2012, pp.503–506.
- [16] Matthews, P. "Social Media, Community Development and Social Capital," *Community Development Journal*, Volume 51, Number 3, 2016, pp.419–435.
- [17] Morgan-Trimmer, S. "It's Who You Know: Community Empowerment through Network Brokers," *Community Development Journal*, Volume 49, Number 3, 2014, pp.458–472.
- [18] Oh, O., Eom, C., and Rao, H. R. "Research Note — Role of Social Media in Social Change: An Analysis of Collective Sense Making During the 2011 Egypt Revolution," *Information Systems Research*, Volume 26, Number 1, 2015, pp.210–223.
- [19] Patino, A., Pitta, D. A., and Quinones, R. "Social Media's Emerging Importance in Market Research," *Journal of Consumer Marketing*, Volume 29, Number 3, 2012, pp.233–237.
- [20] Zeng, D., Chen, H., Lusch, R., and Li, S.-H. "Social Media Analytics and Intelligence," *IEEE Intelligent Systems*, Volume 25, Number 6, 2010, pp.13-16.
- [21] Zhao, Y. *R and Data Mining: Examples and Case Studies*. Academic Press, 2012.

AUTHOR BIOGRAPHIES

Di Shang is Assistant Professor of Information Systems in the Department of Technology, Innovation and Computer Science at Long Island University Brooklyn. His research focuses on information economics, electronic commerce, and business analytics. His work has applied social media analytics techniques to discover insights on community development, disease surveillance and consumer brand awareness.

Mohammed Ghriga is Chair and Associate Professor of Computer Science in the Department of Technology, Innovation and Computer Science at Long Island University Brooklyn. His research interests lie in the broad area of formal description techniques, systems analysis, model checking, testing and verification for non-deterministic systems. He also investigates issues at the intersection of software engineering and security.