Understanding the Behavior of Filipino Twitter Users during Disaster

John Boaz Lee, Michael Ybañez, Marlene M. De Leon, and Ma. Regina E. Estuar

***Abstract*—The Philippines is a country that frequently experiences disasters, such as typhoons. During these events, many citizens spread information and communicate with each other through social media like Twitter. This study aims to take advantage of that fact by analyzing the data from social media to get some insights on the situation. Specifically, this paper studies the behavior of Filipinos on Twitter during a disaster, and tries to see the differences between participants, or the direct victims of the disaster, and observers. The study used Latent Dirichlet Allocation and Principal Component Analysis to extract the different topics discussed during a disaster, and found out which topics participants are more likely to talk about. Results also show which topics are more likely to be retweeted, which language participants in disaster use more often, and what emotions are present in the disaster-time tweets of Filipinos.**

***Index Terms*—disaster, sentiment, social media, social networks, Twitter**

I. Introduction

THE Philippines is a country that is frequently distressed by tropical cyclones and other natural disasters such as earthquakes. Every year, eight to nine tropical cyclones (on average) make landfall in the Philippine Area of Responsibility (PAR) [18]. Earthquakes also hit the Philippines, such as the recent 6.9 magnitude earthquake in the Visayas region [17]. In 2009, 22 tropical cyclones entered and developed inside the PAR, and one of these was typhoon Ondoy (international name: Ketsana), which devastated the Luzon area in September [19] and was quickly followed by typhoon Pepeng (Parma) in early October, aggravating the damage, which was estimated to be 4.3 billion US dollars after these two storms hit [19]. Indeed, the Philippines is a very disaster-prone country.

The Philippines is also a country that is very attuned to social media and mobile technology. The country is nicknamed the “social media capital of the world” [21]. In the Asia-Pacific region, it is actually the country with the highest social networking penetration [10]. It is the eighth most popular country for Twitter[1] use globally [21], and government agencies even have Twitter accounts to spread advisories or warnings to Filipinos, such as the MMDA Twitter account which sends updates on traffic conditions in the Metro Manila area [24].

Disasters and Twitter have close ties with each other. For example, epidemic outbreaks can spur people to talk about their opinions and experiences on Twitter [4]. In fact, studies have been done to analyze whether Twitter is a social network or a news media outlet, and some initial results show that some characteristics of Twitter deviate from other social media, and that the majority of the user-generated content on Twitter is news-related [15]. There have been studies on identifying real world events depicted on Twitter [1], and these studies have been integrated into applications such as real-time earthquake warning and reporting systems [22]. Twitter can contribute various categories of information to situational awareness during hazards [26], and it is this realization that encourages us to further our understanding of Twitter and its users in the Philippines, in order to contribute to disaster response efforts in the Philippines.

In this paper, we study the behavior and characteristics of two different types of Filipino users on Twitter during a natural disaster. This allows us to gain additional insights into the use of social media during times of crisis.

II. Review of Related Literature

Twitter is a microblogging website where users may post short messages of up to 140 characters; each of these posts is referred to as a “tweet”. In order to put some context or topic information into a tweet, users may use “hashtags”, such as #Curiosity or #Olympics2012, into the text of the tweet. Twitter users may subscribe to another user’s updates by “following” that user, who may choose to “follow” back if he or she wishes. Twitter users may choose to subscribe to a “list”, which is a group of users connected together by some shared hobby, interest, or even just by geographical proximity. Users can then use the list feature to view only the tweets by users in that list. Twitter users may also “retweet” another user’s tweet, which is basically copying that tweet to his own Twitter feed.

Several studies have been made to try and discover or filter information from the massive volume of tweets on Twitter as well as its userbase. Yamaguchi et al [29] attempted to tag users based on what topics they are more likely to tweet about using the information from their Twitter lists. Another study ranked users according to their authority or influence on other users, based on how information flows between users [30]. Some

All of the authors are with the Department of Information Systems and Computer Science, Ateneo de Manila University, Philippines.

This paper is based on previous work entitled Characterizing Behavior and Features of Participants and Observers during Disaster on Twitter [12].

[1] http://www.twitter.com

DOI: 10.5176/2251-3043_3.2.254

interesting results also show that even if a Twitter user has many followers, he or she may not necessarily be influential [3]. These studies generally utilize followers-followed relationships and retweets in order to determine authority or influence.

It is also possible to get very useful information about the tweets, and the topics or events discussed in the tweets, rather than the users. Live streams of tweets can be clustered together, and mapped to actual real-world events in real-time [1]. Other approaches to information gathering are more geared towards disaster events: particle filters that consider each Twitter user as a sensor are able to analyze tweets in real time in order to detect whether an earthquake is currently happening [22]. This algorithm was integrated into an earthquake reporting system in Japan that e-mails registered users when the Twitter feed indicates that an earthquake is occurring, with 96% of at least intensity 3 earthquakes being detected correctly [22].

There are many studies that attempt to examine online social networks such as Twitter and Facebook by using graphs to model. The use of graphs may allow researchers to explore other problems and approaches to solving these problems. An example of a graph-based problem is community detection, which attempts to identify groups of nodes that are densely connected to one another and less densely connected to the rest of the network [7], [16], and there are clustering methods used to detect communities in social networks [7]. Another problem that utilizes a graph structure is link prediction. This studies the nodes and the structure of the graph in order to predict the emergence of new links between nodes in the graphs [8], [13]. This is especially useful in recommender systems for social networks that predict which other users a specific user will "befriend." Graphs allow influence propagation in a real world network to be studied: advisor-advisee relationships in a bibliographic network can be identified through influence models [28], and models of topic-level influence of one node on another can predict user behavior for social networks like Twitter and Digg [14]. Clustering has also been studied in the context of networks, and has been used to generate good clustering information in the network [23].

Community detection in social networks is a line of analysis on networks to provide further conclusions. In many networks, the property of community structure dictates that nodes are joined together in tightly knit groups or communities, which are connected to each other, but not as tightly [9]. In a study on collegiate social networks, it was shown that year level and dormitory (or proximity) were highly correlated to the formation of communities in the network [25]. The paper concluded that common residences can generally encourage the formation of new friendships, and vice versa. With regard to the algorithms for doing community detection, one algorithm for detecting communities in social networks is Walktrap, which uses random walks over the network to capture its community structure [20]. This is a new algorithm that is efficient, captures much information, and is of a higher quality compared to previously proposed algorithms for community detection.

The studies most relevant to this paper are those that deal with the usage of social media during disaster. In [11], researchers observed that Twitter messages sent during emergency events seem to be more about relaying information. Vieweg et al. analyzed tweets by people who were "on the ground" during two natural disasters in 2009, and discovered that a substantial number of tweets contributed to situational awareness, which is a "state of understanding 'the big picture' during critical situations" [27]. Another paper describes a set of NLP-based features that can be used by a classifier to identify tweets that contribute to situational awareness.

For the purposes of this study, the authors employed Latent Dirichlet Allocation [2], which is a model for discovering the latent topics that are present in a corpus and to automatically identify the topic distribution for each document in the corpus. It is a generative probabilistic model that can be used for text corpora, of which the dataset in this study is one. In LDA, each item of the collection is given a topic probability for each of different topics generated.

III. DATASET

In order to study behavior in disaster, we gathered a set of tweets about the 2012 flooding in the Philippines caused by the Southwest Monsoon, also known as Habagat. This occurred between August 6–9, 2012. During the event, a script was run using Twitter's streaming API to harvest tweets with keywords related to the event, like "flood" and "Habagat". The entire dataset consisted of around 1.5 million tweets. We gathered the first 5000 tweets from the dataset, and proceeded to manually label each tweet as either a "participant" or "observer"; this number was selected owing to time constraints. A tweet is labeled "participant" if it details first-hand experience of the flooding, or other consequences thereof. For example, a tweet that says "we have to leave again. the flood is almost waist-high inside our house. It reached the outlets; so we need to shut down electricity." is clearly referring to a flood in the immediate vicinity of the tweeter. Tweets about being stuck in traffic (in the context of Habagat) are also considered participant tweets.

We do not consider tweets talking about the rain as tweets by a participant, as the rain covered or affected very large regions, but the flooding affected select areas only; otherwise it would result in a very high number of "participant" tweets that do not provide much information, as many outlying areas could still have experienced a small amount of rainfall. If the tweet is broadcasting information about the flood or related information but only as a news advisory instead of a firsthand account, it will be considered an "observer". In general, if a tweet cannot be labeled as participant, it is labeled an observer (and thus will capture foreign language tweets and other noise in the dataset).

From the sample of 5000, there were 353 tweets by 341 unique participants, and 4647 tweets by 3763 unique observers. Table I shows a small sample of tweets and their labels.

For the purposes of this paper, Twitter users who are the authors of tweets labeled as "participant" and "observer" are themselves called participants and observers, respectively. If a user has more than one tweet from the dataset, he or she is called

TABLE I
LABELS OF SAMPLE TWEETS

Label	Tweet
Participant	Yup; we're going to use this inflatable boat! Wish us luck. Flood is everywhere!!! God bless us Philippines. http://t.co/PzjB8wFJ
Observer	RT @ANCALERTS: PAGASA 4:30pm advisory: RED warning for Metro Manila. Expect heavy to intense rains with occasional torrential rains with ...
Participant	Its raining hard again......flood is rising up again :(
Participant	spent 2 days eating bananas and soup. Flood and rain please go away. I'm running out of supplies here
Observer	RT @rodmagaru: RED WARNING again here in Metro Manila. #FloodPH
Observer	@musamanila think its best to take the MRT. heard traffic's terrible.
Participant	@aimifrances To be honest;Im perfectly fine. Super thankful sa mga household help. Kung wala sila baka nagpalunod nako sa baha. O_O
Observer	Habagat is strong but we've got a much stronger GOD #PrayForThePhilippines #bangonPilipinas #bangonPinoy

a participant if at least one of these tweets is labeled participant.

IV. EXPERIMENTS AND RESULTS

This section outlines the results from the different experiments, organized based on the experiment performed.

A. *Tweet Topics*

To identify the major topics among the tweets in our dataset, we used the LDA topic model [2]. We fed the tweets into the algorithm and specified 8 as the number of topics to be generated; we used this arbitrary number so that a preliminary study could be made. Table II shows the different classifications generated by LDA, as well as the keywords associated with each topic. It is noticeable that some of these results are actually duplicate categories, because of the large number of topics specified. The keywords for each category were inspected in order to label the topic. Three of these topic categories were about inquiries and proclamations regarding class suspensions. One of these topics (topic 4 in the table) was a general catch-all category for conversations between two Twitter users in the Filipino language, as well as their first-hand accounts of their experiences during the disaster. Topics 1 and 2 contained tweets with important service announcements or news updates, for traffic and weather, respectively. Topics 5 and 6 contained keywords regarding rescue and requests.

Figure 1 shows the distribution of topics from tweets of participants and observers. It is interesting to note that the majority of the tweets that came from actual participants (those who experienced the flooding first-hand) were primarily found under topics 4 and 5, which were first-hand accounts and requests for prayer and rescue respectively. Unlike the observers, participants were not very interested in tweeting

TABLE II
TOPICS OF TWEETS DURING DISASTER

Topic	Key Terms
1 (traffic updates)	passable, not, flood, alert, vehicles, deep, types, waist, *[multiple street names follow]* balara, garcia, c.p., mmda, timog, morato, sgt
2 (weather agency updates)	manila, red, warning, metro, signal, rains, pagasa, with, torrential, areas, august, 2012, issued, from, target, heavy, stay
3 (suspension of classes)	aug, antipolo, classes, school, levels, please, public, private, suspended, tomorrow, informed, that, both, @deped ph, mayor
4 (some firsthand accounts, general conversations in Filipino)	*[words in Filipino follow]* ang, lang, yung, naman, ako, mga, dito, hindi, wala, kayo, din, bangonpilipinas, may, walang, pero, lakas, sana, rin, nga, haha
5 (prayer, rescue)	flood, for, you, please, this, that, need, help, relief, edsa, reliefph, prayforthephilippines
6 (relief goods, rescue)	please, help, rescueph, people, water, evangelista, xavierville, don, send, rescue, house, river, you, want, more, goods, heights, evacuees, don't, pls
7 (suspension of classes)	public, classes, makati, private, suspended, tom, aug, mayor, @makatitraffic, orders, college, levels, preschool, on, 2012, aug, advisory
8 (suspension of classes)	san, tomorrow, classes, for, state, calamity, levels, pasig, august, needs, suspended, 2012,

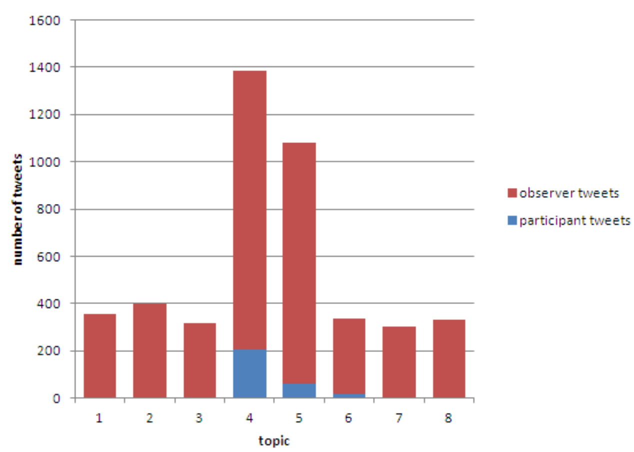

Fig. 1. Topic Distribution of Tweets

about class suspensions, or inquiring about the weather.

Figure 2 is a similar chart; it shows the topic distribution of retweets. It is noticeable that topic 4 tweets, which are about firsthand accounts, are the least retweeted. People were most likely to retweet tweets from topic 5, tweets about prayer and rescue. This can possibly indicate a tendency to retweet "sensational" tweets or tweets from major media operators about rescue operations. This may also be because tweets from participants are sparse and may not contain much useful information in their hurried writing. Participant tweets were not retweeted as much as the observer tweets, and were thus buried beneath more informative retweets.

As an additional experiment, we used Principal Component Analysis to also try to find the initial themes and topics of the tweets. Three stopword dictionaries were created, namely: Twitter stopwords, which comprised standard twitter terms;

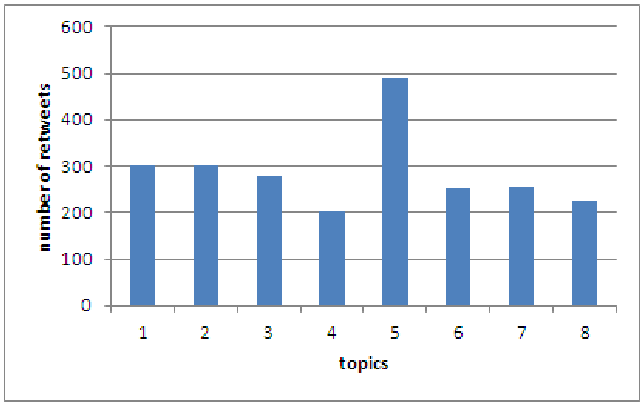

Fig. 2. Topic Distribution of Retweets

Habagat stopwords, which comprised search words used in collecting tweets; and noise stopwords, which comprised single alphanumeric characters including unrecognizable characters. The dataset was pre-processed by transforming all words into lowercase, tokenizing to create the word list, then filtering by using the standard stopword dictionary in addition to the three stopword lists created. The data set was further filtered by removing words that appear in less than 5% of the document. The initial run of PCA showed that there were five possible dimensions with eigenvalues summing up to a cumulative variance of 60.9%. There were five themes that emerged, and these all matched a topic derived from the LDA experiment. The themes are as follows: Effect (of the disaster), Weather, Announcements, Locations, and Requests.

B. *Tweet Language and Length*

The average tweet length for participants was 13.64 words, while average tweet length for observers was 16.71 words. This is probably because participants in the disaster did not have time to be verbose, while many observers retweet advisories from news outlets, which are quite long.

Figure 3 shows the different languages used in the tweets by participants and observers. Participants seem to favor their native language, while observers primarily tweet in English. This may be due to observers retweeting from major media outlets, who usually tweet their advisories in English.

The following additional experiment was performed in order to discover whether the aforementioned result regarding language used was a deviation from the normal behavior of users. The list of unique participants was acquired, and only those who had public profiles were studied. Their tweet history was accessed, and the last 30 tweets from each participant was gathered. These tweets were published after the duration of the Habagat, and were not about the disaster. Figure 4 shows the language distribution of general tweets. In contrast to the disaster-related tweets, where around 75% of participant tweets were in Filipino, around 53% of the general tweets made by the same users were in English.

C. *Tweet Sentiment*

For the purpose of sentiment analysis, all participant tweets were assumed to be subjective. On the other hand, a

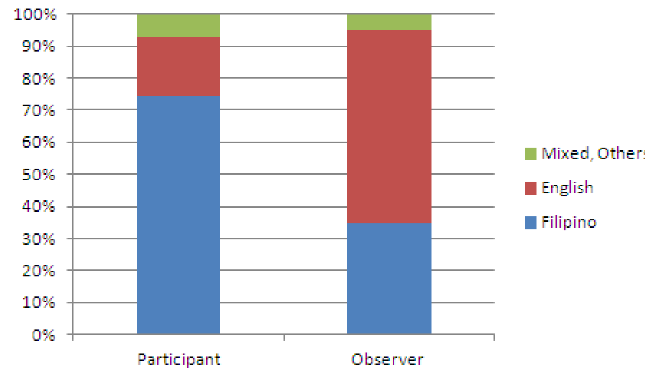

Fig. 3. Language Distribution of Disaster- Related Tweets

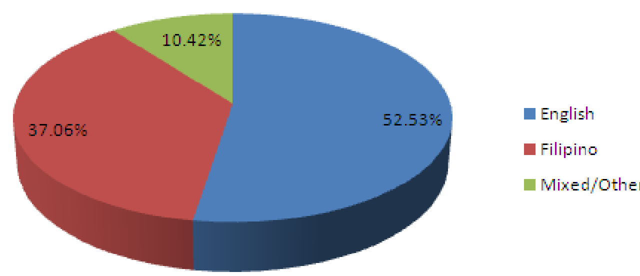

Fig. 4. Language Distribution of General Tweets

subjectivity classification model [5] was used to identify the subjective tweets from all observer tweets. All the subjective tweets were further classified as either positive or negative using a sentiment classification model [6]. Table III shows the sentiments of both participant and observer tweets. Although the majority of tweets are negative in both cases, it is interesting to note that a larger fraction of participant tweets (66%) are negative compared to observer tweets (56%).

The term frequency-inverse document frequency (TF-IDF) vectors for each class of tweets were obtained during sentiment analysis and Principal Component Analysis (PCA) was then applied to the vectors to identify the underlying themes which may be associated with different types of emotions. We applied PCA with a threshold variance of 95%. Table IV displays the various themes that were identified from the various tweets. While the observer tweets possessed a more varied range of emotions, it is interesting to note that Filipinos who are affected by the disaster make light of the situation, and even laugh about it. In terms of negative tweets, participant tweets cover a broader range of emotions and display more extreme emotions like anger.

D. *Communities in the Network*

A network was constructed out of the nodes in the dataset through the following process. First, all the users who tweeted the tweets in the dataset were considered. For clarity, these users will be referred to as the base users. Next, the lists of all the users they followed were then extracted. Finally, for each

TABLE III
SENTIMENT OF PARTICIPANT AND OBSERVER TWEETS

Participants		Observers	
Polarity	Count	Polarity	Count
positive	119	positive	303
negative	234	negative	381
total	353	total	684

TABLE IV
THEMES AND EMOTIONS IN TWEETS

PC	Theme/s	Emotion/s
	Positive observer tweets	
PC1	Pleading for help	Anxiety, weariness, concern
	Thanking for help	Gratitude
PC2	Hoping for something better	Hope
PC3	Laughter despite the situation	Happiness
PC4	Wishing for God's blessing	Hope, sympathy
	Positive participant tweets	
PC1	Laughter and surprise despite the situation	Happiness, surprise
PC2	Laughter despite the situation	Happiness
	Thanking for help	Gratitude
PC3	Full of thanks	Gratitude
	Negative observer tweets	
PC1	Surprise or dismay over the situation	Surprise, distress
PC2	Disbelief over the situation	Surprise, disbelief
PC3	Exasperation and irritation over situations	Exasperation, irritation
PC4	Not liking the situation	Dislike
	Negative participant tweets	
PC1	Pleading to God	Weariness, anxiety
	Exasperation and disbelief over the situation	Disbelief, exasperation
PC2	Expression of surprise	Surprise
PC3	Expression of sadness and disappointment	Sadness, disappointment
PC4	Expression of sadness and anger	Sadness, anger

follower-followed relation where the followed user was not in the set of base users, that relation was dropped. The graph was then constructed from the remaining relations, and was therefore only composed of the edges between the base users, and only those base users who followed at least one other base user. This means that there were no single nodes in the graph not connected to any other node. A total of 2900 nodes existed in the final graph. For simplicity, the graph was undirected.

The Walktrap [20] algorithm implementation in the R software was used on the resultant graph. Table V shows the sizes of the largest communities, divided into participants and observers, with a varying number of steps as the parameter for Walktrap. (The number of steps controls how long the algorithm gathers information about the communities.) Only communities with at least 50 members were included. It is very interesting to note that no communities are composed primarily of participants. It seems counterintuitive that such a result occurs, especially since other studies [25] show that proximity between users has a correlation with these users being in a community structure. It is not difficult for one to assume that participants in a disaster live in areas close to each other; therefore, one should expect that there would be communities where participants are dominant. This could indicate that there are other factors that affected the network and community structure that have not yet been considered, and were not captured in the study. It is possible that, when studying the social network on a global scale, and only considering the

TABLE V
BREAKDOWN OF LARGEST COMMUNITIES

Steps	Participants	Observers	%Participants
2	24	290	8.3%
	101	1312	7.7%
5	62	793	7.8%
	7	67	10.4%
	63	761	8.2%
	16	176	9.1%
10	2	59	3.4%
	54	690	7.8%
	21	242	8.7%
	51	627	8.1%
	8	79	10.1%
	8	81	9.9%
	4	79	5.1%

tweets about the disaster, the Filipino users would comprise a giant community, and one may point to a general geographic location where the disaster is occurring; however, when studying just a subset of the Philippine network, it might be possible that community structure is less useful in separating participants from observers. At a micro-scale, the formation of communities might not be governed by a single dominant factor, such as geographic location.

V. CONCLUSION

Some interesting results were generated. The difference in the topic distribution of the tweets of participants and observers was described. From these topics, tweets regarding requests for prayer and rescue were the most likely to be retweeted. Another observation was in the language use of participants. While most of them have a propensity to use English, they have a larger probability of using their native tongue when tweeting during a disaster. These different observations will help pave the way for further study, such as creating a model to distinguish a participant from an observer. It was also found that community structure in the local social network might not be as helpful in identifying participants and observers. It was also found that direct victims also exhibited emotions such as surprise, happiness, and gratitude. As a recommendation for further study, expanding the data with regards to the nodes and edges included in the network may aid in discovering new insights with regards to the Twitter user's behavior. A study on accurate participant/observer classification is also a promising future direction.

ACKNOWLEDGMENT

The authors would like to thank the Social Computing Science laboratory of the Department of Information Systems and Computer Science (DISCS) for their support, as well as Mr. Hubert Ursua for the Habagat dataset.

References

[1] H. Becker, M. Naaman, and L. Gravano, "Beyond trending topics: Real-world event identification on Twitter," in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 2011.

[2] D. M. Blei, A. Y. Ng, M. I. Jordan, and J. Lafferty, "Latent dirichlet allocation," *Journal of Machine Learning Research*, 2003.

[3] M. Cha, H. Haddadiy, F. Benevenutoz, and K. P. Gummadi, "Measuring user influence in twitter: The million follower fallacy," 2010.

[4] C. Chew and G. Eysenbach, "Pandemics in the age of Twitter: Content analysis of tweets during the 2009 H1N1 outbreak," *PLoS ONE*, 2010.

[5] M. M. De Leon and M. R. E. Estuar, "Deconstructing disaster tweets: Towards building a model for bilingual subjectivity analysis," in *Proceedings of the 8th Engineering Research and Development for Technology Conference*, 2012.

[6] ——, "Disaster emotions: A bilingual sentiment and affect analysis of disaster tweets," in *Proceedings of the 6th Annual International Conference on Computer Games, Multimedia and Allied Technologies, 2013*.

[7] S. Fortunato, "Community detection in graphs," *Physics Reports*, 2010.

[8] L. Getoor and C. P. Diehl, "Link mining: A survey," *SigKDD Explorations Special Issue on Link Mining*, pp. 3–12, 2005.

[9] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[10] GMA News Online, "RP highest in social networking engagement across Asia-Pacific," Apr. 2010. [Online]. Available: http://www.gmanetwork.com/news/story/187883/scitech/rphighest-in-social-networking-engagement-across-asia-pacific

[11] A. Hughes and L. Palen, "Twitter adoption and use in mass convergence and emergency events," in *Proceedings of the 6th Conference on Information Systems for Crisis Response and Management*, 2009.

[12] J. B. Lee and M. Ybañez, "Characterizing behavior and features of participants and observers during disaster on twitter," in *Proceedings of the 6th Annual International Conference on Computer Games, Multimedia and Allied Technologies*, 2013, pp. 76–81.

[13] D. Liben-Nowell and J. Kleinberg, "The link prediction problem for social networks," in *Proceedings of the twelfth international conference on Information and knowledge management*, New York, 2003.

[14] L. Liu, J. Tang, J. Han, M. Jiang, and S. Yang, "Mining topic-level influence in heterogeneous networks," in *Proceedings of the 19th ACM international conference on Information and knowledge management*, New York, 2010, pp. 199–208.

[15] S. Moon, H. Kwak, C. Lee, and H. Park, "What is Twitter, a social network or a news media?" in *Proceedings of the 19th international World Wide Web conference*, New York, 2010, pp. 591–600.

[16] M. E. J. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, pp. 8577–8582, 2006.

[17] M. Palatino, "Philippines and disasters," Feb. 2012. [Online]. Available: http://thediplomat.com/asean-beat/2012/02/09/philippinesand-disasters/

[18] Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA), "Member report to the ESCAP/WMO typhoon committee, 41st session," PAGASA, Tech. Rep., 2009.

[19] ——, "Member report to the ESCAP/WMO typhoon committee, 42st session," PAGASA, Tech. Rep., 2010.

[20] P. Pons and M. Latapy, "Computing communities in large networks using random walks," *J. of Graph Alg. and App. bf*, vol. 10, pp. 284–293, 2004.

[21] J. Russell, "Philippines named social networking capital of the world," May 2011. [Online]. Available: http://asiancorrespondent.com/54475/philippines-named-the-socialnetworking-capital-of-the-world-indonesia-malaysia-amongst-top-10/

[22] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes Twitter users: real-time event detection by social sensors," in *Proceedings of the 19th international World Wide Web conference*, New York, 2010.

[23] Y. Sun, Y. Yu, and J. Han, "Ranking-based clustering of heterogeneous information networks with star network schema," in Proceedings of the 15th *ACM SIGKDD international conference on Knowledge discovery and data mining*, 2009, pp. 797–806.

[24] C. Tonio, "10 must-follow twitter accounts in the philippines," Aug. 2011. [Online]. Available: http://cheftonio.blogspot.com/2011/08/10-must-follow-twitter-accounts-in.html

[25] A. L. Traud, E. D. Kelsic, P. J. Mucha, and M. A. Porter, "Comparing community structure to characteristics in online collegiate social networks," *SIAM Review*, vol. 53, no. 3, pp. 526–543, 2011.

[26] S. Vieweg, A. Hughes, K. Starbird, and L. Palen, "Microblogging during two natural hazards events: what Twitter may contribute to situational awareness," in *Proceedings of the ACM conference on Computer Human Interaction*, New York, 2010.

[27] ——, "Microblogging during two natural hazards events: What twitter may contribute to situational awareness," in *Proceedings of the 28th Conference on Human Factors in Computing Systems*, 2010, pp. 1079–1088.

[28] C. Wang, J. Han, Y. Jia, J. Tang, D. Zhang, Y. Yu, and J. Guo, "Mining advisor-advisee relationships from research publication networks," in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, 2010, pp. 203–212.

[29] Y. Yamaguchi, T. Amagasa, and H. Kitagawa, "Tag-based user topic discovery using Twitter lists," in *Advances in Social Networks Analysis and Mining (ASONAM)*, 2011, pp. 13–20.

[30] Y. Yamaguchi, T. Takahashi, T. Amagasa, and H. Kitagawa, "TURank: Twitter user ranking based on user-tweet graph analysis," in *Proceedings of the 11th international conference on Web information systemsengineering*, Berlin, 2010, pp. 240–253.

John Boaz Lee John Boaz Lee is an Instructor at the Department of Information Systems and Computer Science, Ateneo de Manila University. He received his Master's in Computer Science from the University of the Philippines Diliman.

Michael Ybañez Michael Ybañez is an Instructor at the Department of Information Systems and Computer Science of the Ateneo de Manila University. He gained both his Bachelor's and Master's Degrees in Computer Science at the Ateneo de Manila University.

Marlene de Leon Marlene M. De Leon, Ph.D. studied at the Ateneo de Manila University where she received her Bachelor's in Management Information Systems and Master's and Doctorate in Computer Science. She currently works as a full-time Assistant Professor at the Department of Information Systems and Computer Science (DISCS), where she teaches Software Engineering, Systems Analysis and Design, and Database Management. She currently serves as Associate Chair of DISCS.

Ma. Regina Estuar Ma. Regina E. Estuar, Ph.D. studied at the Ateneo de Manila University where she received her Bachelor's and Master's in Computer Science, and her Doctorate in Psychology. She currently works as a full-time Associate Professor at the Department of Information Systems and Computer Science (DISCS), where she teaches Health Informatics and Research Methods. She heads the Social Computing Science Laboratory and is current the Executive Director of the Ateneo Java Wireless Competency Center.