Investigating Situational Awareness in Microblog Posts during Three Natural Disasters in the Philippines

CELESAMAE T. VICENTE

Abstract

This is an analysis of microblog posts that were broadcasted during three natural disasters in the Philippines namely Typhoon Yolanda, Typhoon Glenda, and Typhoon Ruby. More specifically this research aimed to investigate the major themes of topics and there relevance to the study events and determine the percentage of occurrence of Situational Awareness in the microblog posts when the disasters took place. Techniques used in the analysis include word frequency count, tweet relevance analysis, K-means word clustering, and percentage calculation of relevant updates on Situational Awareness domains.

Keywords: microblogging; situational awareness; typhoon; social computing

---

CELESAMAE T. VICENTE is currently an Assistant Professor at the Department of Information Technology and is the Graduate Program Coordinator of the School of Computer Studies, MSU-IIT. She has a Master's Degree in Computer Science with a research concentration on Eye Tracking from the School of Science and Engineering of Ateneo de Manila University, Quezon City, Philippines, where she was awarded a DOST-ERDT scholarship and internship grants at the National Taiwan Normal University, Taiwan and Narra Institute of Science and Technology, Japan.

Introduction

Technological advancement has made it possible to have speedy access to information that was not available decades ago. Retrieval and relaying of information via web searching and posting or accessing social media sites can be as quick and available to many from the moment a user publishes an information (Vieweg, Hughes, Starbird, and Palen, 2010).

One of the widely adopted means of social media is Microblogging. Through it, information production and retrieval have become more rapid, repetitive, and wide spread (Starbird, Palen, Hughes & Vieweg, 2010). Researchers have considered microblogging as one of the means in communicating during emergency situations such as natural disasters due to its growing pervasiveness, cross-platform accessibility, and speedy communication scheme. It has also been considered as an excellent source of gathering on-the-ground information when a crisis is happening (Palen, Anderson, Mark, Sicker & Grunwald, 2010).

There are a lot of microblogging websites that are available on the internet such as Pinterest, Tumblr.com, Dipity, Plurk, Flattr, Yammer, and Plerb. One of the most popular is Twitter, which is used in this study. It is a platform where users can view online news and use services that allow posting and interacting with messages which are called "tweets" (Kwak, Lee, Park & Moon, 2010). Tweets can be made publicly available and could give information about events such as natural disasters. In particular, when Typhoon Yolanda, Typhoon Glenda, and Typhoon Ruby occurred, Twitter users broadcasted tweets concurrently. All of these three were categorized as strong tropical cyclones that took place in years 2013 and 2014. While Typhoon Glenda and Ruby were recorded to have high sustained winds, the strongest is Typhoon Yolanda that took 7,000 lives (Ofer Merin MD, Yitshak Kreiss,Lin, Pras, & Dagan, 2014; *GMA News Online*, 2014).With the availability of microblog data, the occurrence and degree of situational awareness when these events took place is then worth investigating . More specifically, the focus of this paper is providing answers to the following questions: (1) What are the major themes of the microblog posts broadcasted when the typhoons Yolanda, Glenda, and Ruby happened in the Philippines? (2) What are the percentages of occurrence of Situational Awareness across all domains and whether or not there is an increase or decline of it in comparing the 3 successive major typhoons? Answering these questions will help us investigate the topics that users of Twitter posted, if we can derive major themes from their posts and if these have helped in spreading situational awareness. The succeeding section will further discuss what "Situational Awareness" is, its domains and its importance during Natural Disasters.

Situational Awareness during Natural Disasters

Bedny and Meister (1999) Situational Awareness(SA) refers to:

"the conscious dynamic reflection on the situation by an individual. It provides dynamic orientation to the situation, the opportunity to reflect not only the past, present and future, but the potential features of the situation. The dynamic reflection contains logical-conceptual, imaginative, conscious and unconscious components which enable individuals to develop mental models of external events".

In addition, Sarter defines SA to be "all knowledge that is accessible and can be integrated into a coherent picture, when required to assess and cope with a situation" (Sarter & Woods, 1991). It is achieved by the contribution of information by those who are experiencing a stressful situation through different means, one of which is microblogging (Vieweg, Hughes, Starbird & Palen, 2010).

Microblogged information is now considered by researchers as a source that can be contributory to enhance *SA* of people during safety critical settings. It has become a mean for people to organize information and discuss what they are going through in catastrophes like an earthquake that has struck their community. Twitter and public forums have helped in dispersing relevant information needed by the victims of these calamities and that organizing help for the areas badly hit were made possible whether for providing relief goods, performing search and rescue, or giving counsel have become more accessible resulting to actions quickly done(Liu, Palen, Sutton, Hughes & Vieweg, 2008). Other disaster-related studies showed that microblogging sites were used to provide public information. Examples of which include the Southern California wildfires in 2007 (Shklovski, Palen & Sutton, 2008), the Virginia Tech and Northern Illinois University shootings in year 2007 and 2008 (Palen, Vieweg, Liu& Hughes,2009; Vieweg, Palen, Liu,Hughes & Sutton,2008), and other disaster-related microblog user participation mentioned in the study of Liu et al. during tsunamis and hurricanes in different countries (Sonnenwald & Pierce, 2000).By using social computing analysis techniques, these research works showed that self-organizing behavior of social media users produced accurate results that contributed to advances in actions and official communications in order to bring help to the people or community in need (Vieweg, Hughes, Starbird & Palen, 2010).

The study of Vieweg et al. has used coding categories for SA on-topic tweets during Natural Disasters. These coding categories are used to classify tweets according to subtopics under a general one. In this context, SA has coding categories that include: Warning, Preparatory Activity, Hazard Location, Flood Level, Weather, Wind, Visibility, Road Conditions, Advice for emergency and for information space, Evacuation Information, Volunteer Information, Animal Management, Damage or Injury Reports, and Crisis Aids (Vieweg, Hughes, Starbird & Palen, 2010).These coding categories are used to investigate if such themes exist in the tweets concurrently broadcasted when typhoons Yolanda, Glenda, and Ruby wreak havoc in the Philippines.

The Study Objectives

This study aimed to investigate the relevance of the tweets about the three typhoons, namely Yolanda, Glenda, and Ruby, that happened in the Philippines by investigating the major themes of topics in every event, and whether there is an occurrence of SA in these microblog datasets when these disasters took place. This will give us an overview of the degree of SA of people in terms of Warning, Preparatory Activity, Hazard Location, Flood Level, Weather, Wind, Visibility, Road Conditions, Advice for emergency and for information space, Evacuation Information, Volunteer Information, Animal Management, Damage or Injury Reports, and Crisis Aids when these events took place. The more specific study objectives are outlined below:

1. To gather information on the study events namely: Typhoon Yolanda, Typhoon Glenda, and Typhoon Ruby.
2. To perform data cleaning and tagging of the Tweets in preparation for dataset analysis.
3. To perform data analysis by using the following techniques:
   a. Calculation of Word Frequencies and Relevance Analysis
   b. Clustering of Words
   c. Calculation of Percentage of Situational Awareness

Methodology

The Study Events

The following typhoons are the events that have been used in the investigation process of this research. Communications made publicly on Twitter during the time that these disasters took place are sources for the datasets used in the analysis.

Typhoon Yolanda

Typhoon Yolanda or also known as Typhoon Haiyan, was recorded as one of the strongest tropical cyclones that took a landfall in November of 2013 and leaving at least 7,000 people dead in the Philippines. It started to form on the 3rd of November that year and dissipated on the 11th of the same month. It has been recorded to have the highest 10-minute sustained wind of 230 kilometers per hour and a 1-minute sustained wind of 315 kilometers per hour. The devastated areas are Leyte and parts of the Eastern Visayas. Storm surges were also recorded with waves estimated to have risen up to 15 feet high leaving more people dead and homeless. Public Storm Warning Signal (PSWS) was initially released, to 80 provinces including the capital Metro Manila, by the Philippine Atmospheric, Geophysical and Astronomical Services Administration (PAGASA) to be of the number 1 signal which is the lowest but was raised signal number 4, the highest, as the storm entered in the Philippine Area of Responsibility. Despite the public warnings that has been issued by the government, Storm Surges occurred and claimed so many lives and others are still missing, especially from those living in the coastal areas (Ofer Merin, et.al, 2014).

Typhoon Glenda

Typhoon Glenda or also known as Typhoon Rammasun was considered to be a powerful tropical storm that entered the Philippine Area of Responsibility on July 10th of 2014 until July 20th of same year. Its highest 10-minute sustained winds were recorded to be 165 kilometers per hour and a 1-minute sustained wind of 250 kilometers per hour. Fatalities were recorded to be at a total of 195. It was the first typhoon to impact the Philippines 8 months after Typhoon Yolanda. Early on the PSWS was already raised to signal number 3, classes were suspended early on, and that the public were already warned for possible landslides, flashfloods, strong rains and winds. Evacuation of residents was also performed due to the threats of storm surges which was estimated to be 4 feet to about 10 feet high. Officials within the areas anticipated to have landfalls of the storm were alert and were preparing for the worst thus, evacuation procedures were done early on (*GMA News Online, 2014*).

Typhoon Ruby

Typhoon Ruby or also known as Typhoon Hagupit formed on December 1st of 2014 and dissipated on the 12th of December same year. It was recorded to have the highest 10-minute sustained winds of 215 kilometers per hour and a 1-minute sustained wind of 285 kilometers per hour. It was recorded to have 18 fatalities. It is recorded to be the second most intense tropical cyclone in 2014. It was considered initially to be of the worst threat that would fall in the Philippines and anticipated to incur damages more than that of Typhoon Yolanda. Fortunately, it was significantly smaller than the other.

Public warnings have indicated Typhoon Ruby to be under the category 5 super typhoon. PSWS was at signal number 1 and 2 in lower parts of Luzon and Northern parts of Mindanao. Class suspensions were also done early on in areas of Samar, Biliran, and Typhoon. Storm surge warnings were up to 4 meters high and residents took precautionary measures against this warning (*GMA News Online, 2014*).

The Data Description, Collection, Cleaning, and Tagging

Analysis was performed on Twitter posts, that have been collected using the Streaming Application Program Interface(API) of Twitter, when the three Typhoons hit the Philippines namely: Typhoon Yolanda in November of 2013, Typhoon Glenda in July of 2014, and Typhoon Ruby in December of 2014. The data gathering process was done by the Social Computing Laboratory of the Ateneo de Manila University wherein they have dedicated computers that collected tweets between years 2013 to 2014. They have already pre-processed the data and grouped them by events that they were and are to conduct investigation on. The datasets needed for this research were already stored in Microsoft Excel - Comma Separated Value(CSV) file format when the researcher performed the cleaning and tagging of tweets. Dataset for each study event was stored on a separate CSV file. These CSV files contained columns for the time, location, and the text content of the tweet. The time and range of datasets that were included were from a week before, during, and after the study events occurred. Geo-location information or the places where the Tweets originated have been disregarded in the analysis of the data. Both English and Tagalog tweets were collected and disregarding all other tweets written in other dialects or foreign languages.

Data cleaning was done by utilizing only text content of the tweets and disregarding all other data columns of the CSV files. Further data cleaning was performed in order to disregard all other tweets that do not mention any related information to the study events. A total of 10,701 tweets have been retained for all study events: 3,297 during Typhoon Yolanda, 4,336 during Typhoon Glenda, and 3,068 during Typhoon Ruby. One by one, data tagging was performed by the researcher in order to categorize the tweets into the SA coding categories: Warning, Preparatory Activity, Hazard Location, Flood Level, Weather, Wind, Visibility, Road Conditions, Advice for emergency and information space, Evacuation Information, Volunteer Information, Animal Management, Damage or Injury Reports, and Crisis Aids as being used in (Vieweg, Hughes, Starbird, & Palen 2010).

The Analysis Tool

Rapid Miner is a software platform that has features wherein users can perform data preparation, machine learning, deep learning, predictive analytics, and text mining which is the main analysis used in this study (Hofmann& Klinkenberg, 2013). It is a data mining tool which has a graphically integrated environment. Users can perform visual component placement, connection, and dragging in it. Once the data has been pulled in the environment it can then be transformed and operations can then be pulled in the visual component. How these operations are implemented are abstracted from the users as operations are already available in a package format. Users just drag and drop operators and run the process they have designed and output based on the operators are then being displayed (Jovic, Brkic & Bogunovic, 2014). In this research, Rapid Miner was used to analyze data using the available features in the form of operators. The detailed list of operators used and their functionalities are discussed in the next section.

The Operators Used for the Analysis Tasks

In the attempt to quantify SA of people posting tweets during the occurrence of the study events, the succeeding techniques were done in

Rapid Miner using the operators available in its software package. The detailed lists of operators that have been used and dragged into the visual component area are listed for each analysis task. At this point, data cleaning has already been performed and that only relevant tweets have been included and that delimiters such as punctuation marks have been removed.

i. *Word Frequencies and Relevance Analysis* – this is to count the occurrence of words and rank them from the highest count to the lowest. This is performed in Rapid Miner by first setting the data source which is in Comma Separated Values (CSV) format. The Nominal to Text operator was used in order to activate text mining. The Process Document operator was then selected in order to set the data source. The Tokenize operator was chosen so that the documents were split into sequence of tokens. In order for several unwanted words to be removed and transformed, operators such as Filter Stopwords for English (removes common English words liked "a", "and", and etc.) and the Filter Stopwords for Dictionary (drop certain words that were not present in the dictionary) was activated. The Total Occurrences and Document Occurrences of a specific word were extracted after running the Rapid Miner. Analysis of relevance of tweets with the study events was then performed.

ii. *Word Clustering* – is done to group words according to major topics in the datasets. Operators used in Rapid Miner were the Process Document Operator, Data to Documents, Tokenize, Filter Stopwords (English), Transform Cases were also used in this process. Other additional operator used was the Stem(Porter), which mapped the relatedness of the words. K-Means Clustering Operator was then used to determine the 4 major themes of topic in the study events. This Rapid Miner operator uses K-means clustering algorithm that classifies the points in a data set into a pre-set fixed number of clusters. It is then performed by setting a central point, or the centroid for each cluster. It is an advantage if the chosen centroids are being set to be far away from each other. Dataset points are then being associated to the nearest centroids and then classified to be part of that cluster. Finding the centroids for the clusters and calculating the distances of the points are calculated using Eucledian distance measurement. Several iterations of centroid assignment is being done when a new dataset point is added (Jovic, Brkic & Bogunovic, 2014). In the context of this study, 4 clusters were set for the dataset of each Typhoon. Although, how the Rapidminer tool runs the K-means Clustering Operator is not observable, the underlying calculations performed were based on how a K-means algorithm behaves. Words were converted to vectors in order to get centroids and measure the relatedness of other words to these centroids.

iii. *Percentage of Situational Updates* – *serves* to quantify the occurrence of words relevant to the domains indicating Situational awareness namely: (a)Warning, (b) Preparatory Activity, (c) Hazard Location, (d)Flood Level, (e)Weather, (f)Wind, (g)Visibility, (h) Road Conditions, (i) Advice for emergency and information space, (k)Evacuation Information, (l) Volunteer Information, (m) Animal Management, (n) Damage or Injury Reports, and (o) Crisis aids. The words tagged for situational updates were the once with the highest relevance to the events. In this particular study, tagging was based on the words with highest frequencies.

Results and Discussions

Word Frequencies and Relevance Analysis

Table 1 shows the result of word frequency analysis done on the three study events namely: Typhoon Yolanda, Typhoon Glenda, and Typhoon Ruby. Only the top 20 words were extracted and shown in the table since all other words were of no relevance to the study events. The top 20 words for Yolanda tweets appear to be precautionary in nature and were giving advices to stay safe and be updated by forecasted news. The Glenda tweets on the other hand were also precautionary in nature, to stay safe, and have also given information about class suspension and constantly referencing PAGASA updates. Interestingly, Ruby tweets were also precautionary in nature, and have referenced advice from the Department of Education(DepEd)and information from news. Ruby tweets also show the intention or need of "aid" and "relief" in some affected areas. In addition, Figure 1 also shows the chart of the number of count for the words with the highest frequency. Percentage relevance of each column of words according to each event was done by calculating the total number of occurrence per word per event and dividing them by the total number of tweets. The highest percentage of relevance was found to be 98.86 % which belongs to the data gathered during Typhoon Glenda, followed by Ruby which is 51.62 % and then by Yolanda which is 38.96 %. These percentages suggest that tweets gathered during Glenda had the highest number of relevance to the situation compared to the other datasets. The Glenda dataset has more relevant information about the study event than the Ruby dataset and the least relevant of which is the Yolanda dataset. The next analysis was done in order to further analyze why the Glenda tweets scored the highest relevance in comparison to the other two study events.

Table 1: Count of Document Occurrence per Word

YOLANDA 2013		GLENDA - July 2014		RUBY - December 2014	
WORD	COUNT	WORD	COUNT	WORD	COUNT
bagyo	596	ulan	982	bagyo	1020
typhoon	566	bagyo	842	typhoon	587
alam	168	weather	574	affected	129
safe	106	signal	481	walangpasok	113
signal	45	typhoon	403	pagasa	101
storm	44	lakas	281	deped	90
news	41	safe	255	laut	67
stay	39	panahon	185	news	57
malakas	38	hangin	152	ulan	40
supertyphoon	37	everyone	136	help	44
ready	35	malakas	125	nationwide	41
visayas	32	ingat	80	victims	39
safety	30	suspended	71	safe	37
paparating	28	classes	67	weather	35
hope	26	pray	61	damage	32
landfall	26	storm	59	donate	32
update	25	flood	55	hangin	30
pagasa	23	pagasa	47	lakas	30
praying	22	cold	48	aid	29
alert	21	home	39	relief	28
TOTAL	1948		4943		2581
PERECENTAGE	38.96%		98.86%		51.62%

NO OF TWEETS
5000

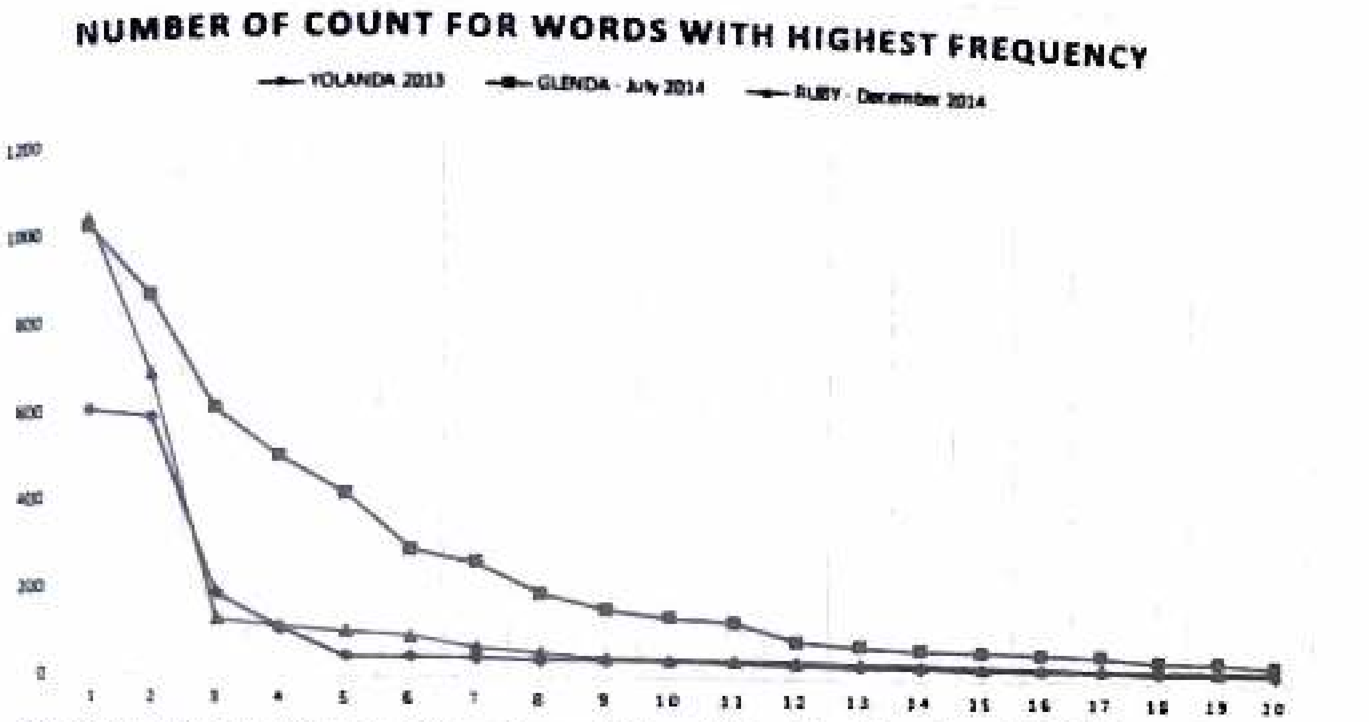

Figure 1: Chart for the Number of Count for the Words with Highest Frequency

Word Clustering

The following subsections discuss the results of the k-means clustering process done in Rapid Miner for each study event. Investigation on themes discussed by cluster in every study event was done. Iterations on the number of clusters were experimented on before deciding that 4 clusters will be used for each study event. This number of major themes was found to best represent the themes discussed on each study event.

Yolanda Clustering

Table 2 shows the clustering results for Typhoon Yolanda tweets. In total there were 3,297 tweets relevant to this study event and running the k-means clustering returned the following distribution: Cluster 0 has 2,613 tweets, Cluster 1 has 392 tweets, Cluster 2 has 253 tweets, and Cluster 3 has 39 tweets.

Table 2. Typhoon Yolanda Clustering Results

Cluster	Data Points Count
0	2,613
1	392
2	253
3	39
TOTAL	3,297

The following are the description of the nature of each cluster:

Cluster 0 includes the words "alam", "bagyo", "typhoon", "supertyphoon", "news", "ready", "storm", "rain", "safety", and "malakas". This cluster indicates a group of people tweeting about warnings that the coming typhoon is strong and that people should stay in safety.

Cluster 1 includes the words "typhoon", "safe", "stay", "hope", "safety", "strong", "praying", "country", and "spare". This cluster indicates a group of people seemingly hoping and praying that the country be spared from the typhoon.

Cluster 2 includes the words "bagyo", "malakas", "paparating", "nakakatakot", "pasok", "parating", "ramdam", "safe", "matuloy", and "lumihis". This cluster shows a group of people anticipating and being fearful at the same time about the coming typhoon.

Cluster 3 includes the words "signal", "areas", "pagasa", "leyte", "eastern", "weather", "southern", "landfall", "storm", "city", and "update". This cluster indicates a group of people tweeting updates about areas that might have been hit by the typhoon.

Typhoon Yolanda Clustering show that about 79% of the tweets were concentrated on warnings about the Typhoon as indicated in Cluster 0. This is indicative that most of the tweets invoke Situational Awareness on the Warning and Preparatory Activities domain before the study event happened.

Glenda Clustering

Table 3 shows the clustering results for Typhoon Glenda tweets. In total there were 4,336 tweets relevant to this study event and running the k-means clustering returned the following distribution: Cluster 0 has 71 tweets, Cluster 1 has 123 tweets, Cluster 2 has 3992 tweets, and Cluster 3 has 150 tweets.

Table 3. Typhoon Glenda Clustering Results

Cluster	Data Points Count
0	71
1	125
2	3,992
3	130
TOTAL	4,336

The following are the description of the nature of each cluster.

Cluster 0 includes the words "signal", "safe", "everyone", and "ingat". This cluster indicates a group of people sending out messages to stay safe and at the same time giving information of the signal number level of the typhoon.

Cluster 1 includes the words "weather", "typhoon", "forecast", and "pagasa". This group of tweets provides information about weather forecasts most likely by PAGASA.

Cluster 2 includes the words "home", "walangpasok", "ramdam", "lamig", and "raining". This cluster indicates group of people sharing the temperature level they experience while they are at home.

Cluster 3 includes the words "ulan", "lakas", "hangin", "bagyo", "safe", "walangpasko", "nakakatakot", "malakas", and "ingat". This cluster belongs to group of twitters who feel the fear and the need to stay in a safe place due to strong winds they are experiencing.

Getting the highest percentage of number of tweets that belong to Cluster which is Cluster 2, 92% of the tweets were concentrated on Weather updates. This is indicative that most of the tweets during this study event invoke the Situational Awareness on the Weather Domain.

Ruby Clustering

Table 4 shows the clustering results for Typhoon Ruby tweets. In total there were 3,068 tweets relevant to this study event and running the k-means clustering returned the following distribution: Cluster 0 has 2,817 tweets, Cluster 1 has 102 tweets, Cluster 2 has 77 tweets, and Cluster 3 has 72 tweets.

Table 4. Typhoon Ruby Clustering Results

Cluster	Data Points Count
0	2,817
1	102
2	77
3	72
TOTAL	3,068

The following are the description of the nature of each cluster:

Cluster 0 includes the words "bagyo", "typhoon", "pagasa", "laut", "news", "deped", "ulan", "lakas", "safe", and "affected". This cluster indicates a group of people tweeting information based from PAGASA and Department of Education (deped).

Cluster 1 includes the words "walangpasok", "nationwide", "deped", "clothes", "lost", "homes", and "donate". This cluster mainly represent a group of people saying that many homes were lost and that donation is needed.

Cluster 2 includes the words "affected", "typhoon", "volunteering", and "help". This cluster of twitters encourages volunteering because many have been affected by the typhoon and needed help.

Cluster 3 includes the word "bagyo". This cluster of people shared information that the focus is on the storm or "bagyo".

Ninety-one percent (91%) of the tweets during this study event belongs to Cluster 0, which invokes Situational Awareness on the Domain of Weather and Advice. Further analysis on Percentage distribution per SA categories will be discussed in the succeeding section.

Percentage of Situational Updates

To know more details on whether or not tweets contribute to situational awareness, words with the highest frequency count per event were tagged according to the following categories discussed in the Analysis Tools and Operators Sections of the Methodology. Grouping the datasets after the data tagging showed that only the following coding categories or domains are present in all of the study events namely: *Warning, Preparatory Activity, Weather, Advice, and Crisis Aids.* The percentage distribution of each category in every study event is shown in Figure 2. Results by study event in Tables 4 and 5based on the number of

SA categories covered showed that Ruby dataset returned the highest for all coding categories. While Yolanda and Glenda lack tweets that mentioned updates on Crisis Aids. However, though there were more number of tweets during Glenda relevant to situational awareness in comparison to that of Yolanda and Ruby datasets, nonetheless, it did not cover as much categories as the Ruby dataset.

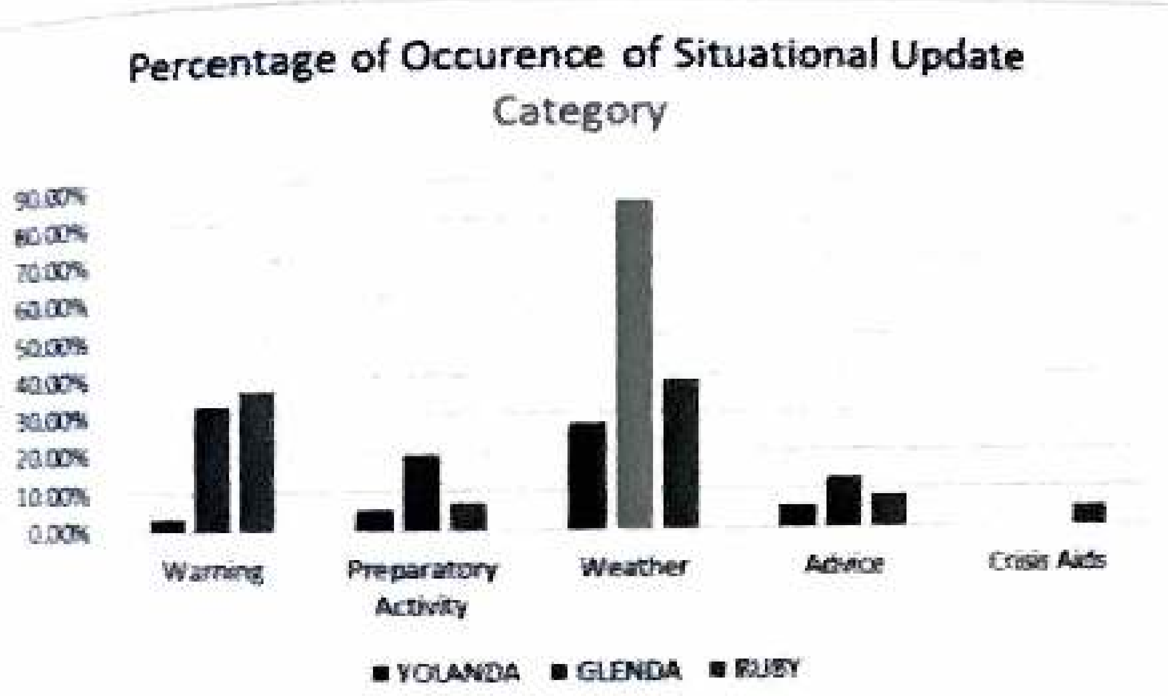

Figure 2. Chart for Percentage of Occurrence of Situational Update Category

Table 5: Percentage of Occurrence of Situational Update Category Per Event

Situational Updates On	YOLANDA	GLENDA	RUBY
Warning	3.54%	32.80%	36.22%
Preparatory Activity	5.56%	19.64%	7.30%
Weather	27.24%	84.66%	38.62%
Advice	5.70%	12.80%	8.50%
Crisis Aids			4.84%

Table 6: Count of Word Occurrence Relevant to the Situational Updates

Situational Updates On	YOLANDA	GLENDA	RUBY
Warning	177	1640	1811
Preparatory Activity	278	982	365
Weather	1362	4233	1931
Advice	285	640	425
Crisis Aids			242
Total	2102	7495	4774

Conclusions

Results show that users contributed more to situational awareness during the Typhoon Glenda but has decreased during the occurrence of Typhoon Ruby in terms of number of tweets posted relevant to the event (See Figure 2). From the nature of the Typhoon Glenda's dataset, micro bloggers have become more precautionary. This might be attributed into seeing the after effects of Typhoon Yolanda. Micro bloggers tend to be more situationally aware in almost all domains or coding categories present depending upon the degree of damage that a previous typhoon has brought. Data sets collected during the Typhoon Glenda in particular have the highest total count of situational awareness related tweets than that of the two other typhoons since it has occurred after that of Yolanda which was documented to have a large number of victims and damage of properties as it was reiterated on different media, including social media sites (Ofer Merin, Yitshak Kreiss, Lin, Pras & Dagan, 2014). Interestingly, situational updates on Crisis Aids have occurred only during Typhoon Ruby; while during Typhoon Yolanda and Glenda, the situational update categories covered were only Warning, Preparatory Activity, Weather, and Advice. This indicates that microbloggers have become more involved in sharing information on helping out the community as past disasters have been experienced and they became more aware of the need to be involved in this domain. This occurrence of specific situational awareness domain may have been contributory into a widespread encouragement for more social media users and viewers to help out the affected individuals to cope with the actual situation. It is also important to note that during the dataset collection for Typhoon Yolanda, people are not microblogging relevant topics to situational awareness as much as that of the two Typhoons. This is because though they know that a typhoon was coming, analyzing unfiltered sets showed that another political topic was trending and was discussed alongside the coming typhoon.

Recommendations

This paper can serve as baseline study for all other detailed investigations that can be done in the datasets. This endeavor is a preliminary effort on getting the general landscape of the data in order to know that further investigations can be done in the context of Situational Awareness and its domains. In addition, to further validate and investigate whether users of Twitter contribute to situational awareness in the natural disasters that occur in the Philippines, it is recommended that a larger dataset maybe investigated and do a more coarse-grained study and expand the network of people who will do the tagging or labeling of tweets according to the situational updates they imply. It is also good to investigate the concentration and origin of the tweets using the geo-tag feature of Twitter. In addition, although this research showed the presence of situational awareness in the analyzed data, it should also be investigated if this has translated to actual mitigation of casualties and other mishap when the typhoons occurred.

Acknowledgement

The author would like to acknowledge the Ateneo De Manila University Social Computing Laboratory for gathering and graciously sharing the data needed and used in the analysis. Especially to Doctor Ma. Regina E. Estuar and Mr. Noel Victorino. Glory be to God!

References

Bedny, G., & Meister, D. (1999). *Theory of activity and situation awareness. International Journal of cognitive ergonomics*, 3(1), 63-72.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1), 100-108.

Hofmann, M., & Klinkenberg, R. (Eds.). (2013). *RapidMiner: Data mining use cases and business analytics applications*. CRC Press.

Jovic, A., Brkic, K., & Bogunovic, N. (2014, May). An overview of free software tools for general data mining. In *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention on* (pp. 1112-1117). IEEE.

Kwak, H., Lee, C., Park, H., & Moon, S. (2010, April). What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web* (pp. 591-600). ACM.

Liu, S. B., Palen, L., Sutton, J., Hughes, A., & Vieweg, S. (2008, May). In search of the bigger picture: The emergent role of on-line photo sharing in times of disaster. In *Proceedings of the information systems for crisis response and management conference (ISCRAM)* (pp. 969-980).

Ofer Merin MD, M. H. A., Yitshak Kreiss, M. D., Lin, G., Pras, E., & Dagan, D. (2014). Collaboration in response to disaster--Typhoon Yolanda and an integrative model. *The New England journal of medicine, 370*(13), 1183.

Palen, L., Anderson, K. M., Mark, G., Martin, J., Sicker, D., Palmer, M., & Grunwald, D. (2010, April). A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters. In *Proceedings of the 2010 ACM-BCS visions of computer science conference* (p. 8). British Computer Society.

Palen, L., Vieweg, S., Liu, S. B., & Hughes, A. L. (2009). Crisis in a networked world: Features of computer-mediated communication in the April 16, 2007, Virginia Tech event. *Social Science Computer Review*, 27(4), 467-480.

Sarter, N. B., & Woods, D. D. (1991). Situation awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology*, 1(1), 45-57.

Shklovski, I., Palen, L., & Sutton, J. (2008, November). Finding community through information and communication technology in disaster response. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 127-136). ACM.

Sonnenwald, D. H., & Pierce, L. G. (2000). Information behavior in dynamic group work contexts: interwoven situational awareness, dense social networks and contested collaboration in command and control. *Information Processing & Management*, 36(3), 461-479.

Starbird, K., Palen, L., Hughes, A. L., & Vieweg, S. (2010, February). Chatter on the red: what hazards threat reveals about the social life of microblogged information. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 241-250). ACM.

Typhoon Hagupit | GMA News Online | URL: http://www.gmanetwork.com/news/news/nation/391145/typhoon-ruby-hagupit-person-finder-and-crisis-map/story/ | Date Published: December 5, 2014

Typhoon Rammasun | GMA News Online | URL: http://www.gmanetwork.com/news/news/nation/370738/glenda-death-toll-now-at-40-damage-to-crops-at-p2-3b/story/ | Date Published: July 17, 2014

Vieweg, S., Hughes, A. L., Starbird, K., & Palen, L. (2010, April). Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1079-1088). ACM.

Vieweg, S., Palen, L., Liu, S. B., Hughes, A. L., & Sutton, J. N. (2008). *Collective intelligence in disaster: Examination of the phenomenon in the aftermath of the 2007 Virginia Tech shooting*. University of Colorado.