Examining The Impact of Twitter (Tweets) On Security Using Sentiment Analysis and Opinion Mining: A Case Study of Nigeria

Salisu, S & Longe, O.B
School of IT & Computing
American University of Nigeria
Yola, Nigeria
E-mail: Shafiu.Salisu@aun.edu.ng; Olumide.longe@aun.edu.ng
Phone: +2348034932870

ABSTRACT

As social media networks continue to grow and increase in popularity, they also create both positive and negative experiences in diverse socio cultural settings. Sentiment analysis and opinion mining have been studied by previous researchers in the context of religion, politics, and hate speech (Ali, 2017; Bauwelinck & Lefever, 2019; Martins & Gomes, 2018). These extant studies have confirmed that sentiment analysis is a vital tool for measuring user generated opinion on twitter, in order to predict or force a change in the community. The thrust of our research is to engage sentiment analysis and data mining to examine the impact of user generated opinions on twitter relatibe to security challenges in socio-cultural settings like Nigeria.

Keywords: Organisation culture, Job performance, Librarians, University libraries in southwestern Nigeria

23rd iSTEAMS Conference Proceedings Reference Format

Salisu, S & Longe, O.B (2020): Examining The Impact of Twitter (Tweets) On Security Using Sentiment Analysis and Opinion Mining: A Case Study of Nigeria. Proceedings of the 23rd iSTEAMS Conference, American University of Nigeria, Yola. April, 2020. Pp 151-156 www.isteams.net/yola2020

1. INTRODUCTION

The new media has created a new norm in the way we think, process, and disseminate information. The new media enabled by social media networks like Facebook, Twitter, and Instagram have changed the way we share sentiment and opinions (Cambero, 2016; Yuan, 2016). Social media networks are generating millions of user opinions on a daily basis, accompanied with wide ranging topics, arising from trending events (Yuan, 2016). These massive chunk of data continues to disrupt the previous normal, causing political, socio cultural change across different continents. For example, the microblogging site twitter has become a very popular go to site for trending events.

Tang and Dalzell (2019) presented a novel idea using two layer model for classifying hate speech on social media; also, in 2010 when the massive 7.0 magnitude earthquake hit mainland Haiti, the news of the events and the subsequent campaign for relief efforts occurred on Twitter. Political transitions around the world have been actively shaped by twitter campaigners, political activist and gladiators; in both the 2008 and 2016 presidential elections in the United States of America Twitter and Facebook were an integral part of the campaign process, and they ultimately shaped the minds and opinions of voters in some quarters (Yuan, 2016). Social media through its enablers (social media networks), is steadily becoming an industry of its own (Subramaniyaswamy, Logesh, Abejith, Umasankar, & Umamakeswari, 2017).

Many companies and industries are taking advantage of this growing niche, in order to understand the opinions of their customers, and socio cultural disruptions. However, this positive development comes with its own challenges such as security and privacy of information. It to this end that social networking companies provide a backend Application Programming Interface (API) to assist developers develop useful tools and models that could help track, analyze and monitor diverse kinds of opinion on social media (Subramaniyaswamy et al., 2017).Enterprises are now faced with diverse security threats from social media, as a result of user generated opinions that are being used by cybercriminals to defraud unsuspecting individuals (Subramaniyaswamy et al., 2017) Therefore, this study will apply a lexicon based sentiment analysis model to twitter data in order to mine useful information. This information can provide insights, detect threats, predict user behavior, thus help organizations and law enforcement agencies make better decisions.

1.1 Problem Statement

In the last half decade, we have witnessed quite a number of security challenges and radical political change fueled by massive social media campaigns and activism. Some of these massive movements in social media have led to political change across the world (Barclay, Pichandy, Venkat, & Sudhakaran, 2014); Arab spring crises (Abuzaakouk, 2015); the earthquake in Haiti (Tang & Dalzell, 2019); and the Bring Back Our Girls movement in Nigeria.

As social media networks continue to grow and increase in popularity, they also create both positive and negative experiences in diverse socio cultural settings. Sentiment analysis and opinion mining have been studied by previous researchers in the context of religion, politics, and hate speech (Ali, 2017; Bauwelinck & Lefever, 2019; Martins & Gomes, 2018). These extant studies have confirmed that sentiment analysis is a vital tool for measuring user generated opinion on social media, in order to predict or force a change in the community.

However, there are paucity of research that contextualize social media based security threats to organizational infrastructures in Nigeria. This work is premised on this assertion and therefore, justifies the need to study the impact of user generated opinions on enterprise information systems in a socio cultural settings such as Nigeria.

1.2 Research Questions

Based on the foregoing, the research question emanates is:

- To what extent can opinion mining assist us to achieve socio-cultural sentiment mining and understanding of security threats to organizational infrastructure?

1.3 Aims & Objectives

1.3.1 Aims

The aim of this research is to assess social media induced threats on enterprise information systems using sentiment analysis, along with the following objectives

1.3.2 Objectives

To achieve the aim of the research, the following objective will be pursued:

i. Employ sentiment analysis techniques to investigate how user opinions and online behavior pose a threat to enterprise information systems.
ii. Develop a model for measuring security threats on social media.
iii. Test the developed model for its efficacy in addressing social media based security threats to enterprise information systems.

2. OVERVIEW OF RELATED WORK

What follows is a summarized view of related literature

Table 1: Table of of Related Literature

AUTHOR	TITLE	METHOD	FINDINGS
(Subramaniyaswamy et al., 2017)	"Sentiment Analysis of Tweets For Estimating Criticality and Security of Events"	"Lexicon Based Model"	"The system they developed combines a lexicon based sentimental analysis along with deep data collection and segregate the emotions into different levels to analyze the threat for an event".
(Alqarni H, Almurtadha Y, and Elfaki A 2018)	"A Twitter Sentiment Analysis Model for Measuring Security and Educational Challenges: A Case Study in Saudi Arabia"	Sentiment Analysis Model	"The model developed in this study was able to predict and categorize tweets belonging to the cross-cultural and ethics of dialogue and rules of difference challenges than those for the dominant negative social value challenge".
(Neuendorf K. A 2017)	The Content Analysis Guidebook	-	Sentiment analysis was used to understand customer behavior.
(Bolle J, Mao H, and Pepe A. 2011)	"Modelling Public Mood and Emotion: Twitter sentiment and socio-economic phenomenon"	Sentiment Analysis	"Profile of Mood States (POMS) was used to model public mood and emotion using twitter sentiment analysis and the outcome was that social, economic and political events do have a direct impact on public mood".
(Al-Khalifa H.S., 2012)	"A First Step Towards Understanding Saudi Political Activities on Twitter"	Graph Analysis	"Findings showed that Saudi Arabia has some underlying social structures in their political landscape"
(Ullmann S., 2017)	"Conceptualizing force in the context of the Arab Revolutions: A comparative analysis of international mass media reports and Twitter posts"	-	Twitter post were a mobilizing for the Egyptian revolution.
(Alabbas W., Al-Khateeb H.M, Mansour A, Epiphaniou G, and Frommholz I., 2017)	"Classification of colloquial Arabic tweets in real-time to detect high risk floods".	Sentiment Analysis	Twitter post were mined to detect high risk floods.
(Desai T., Shariff A., Shariff M., and Fan X., 2012)	"An in-depth analysis of twitter activity at kidney week"	Sentiment Analysis	Results showed that twitter is an effective medium for improving medical care awareness.

AUTHOR	TITLE	METHOD	FINDINGS
(Bauwelinck N., and Lefever E., 2019)	Measuring the Impact of Sentiment for Hate Speech Detection on Twitter	Supervised Classification method with Support Vector Machines (SVMs)	The model developed was able to achieve a high degree of accuracy in detecting hate speech on twitter
(Glass K., and Colbaugh R, 2012)	"Web Analytics for Security Informatics"	-	They developed a system for detecting cyber incidents
(Basave A.E.C., He Y., Liu K., and Zhao J. 2013)	"A weekly supervised Bayesian Model for Violence Detection in Social Media"	Violence Detection Model	The system was able to detect violent related words.
(Colbaugh R., and Glass K., 2011)	"Agile Sentiment Analysis of Social Media Content for Security Informatics Applications".	Lexicon based model	"They system developed was able to estimate regional public opinion regarding the Egyptian revolution and the Jakarta bombing".
(Ngoc P.T., and Yoo M., 2014)	"A lexicon-based sentiment analysis for fan page ranking in Facebook"	-	Their system was able to detect comment polarity and user engagement.

3. METHODOLOGY

This study will adopt the positivistic approach as it tries to find meaning from actual facts emanating from social phenomenon. The study will adopt the quantitative research methodology, along with a lexicon based model. To achieve our objective number 1 and 2 a sentiment scoring technique will be used to score all the data (tweets), the data will be scored between 0-5. Scores above 4 indicates a high security threat, scores less than 3 indicates low security threat. To achieve our objective number 3, the developed model will then be used to analyzed tweets related to selected enterprises that will be used for this study in order to test the efficacy of the model.

4. DATA COLLECTION STRATEGY

The study data will be primarily twitter feeds (tweets) collected from twitter application using twitter developer backend application interface (API). In order to achieve our objectives, the dataset from tweets will span from 2014 – 2019 specifically across the twitter handles of selected Nigerian enterprises.

5. DATA ANALYSIS/DATA PRESENTATION

Data analysis in quantitative research encompasses the use of statistical tools/methods to analyze data in order to bring out useful information (Amaratunga, Baldry, Sarshar, & Newton, 2013). R studio will be used for the data analysis. Descriptions such as graphs, charts, and tables will be used to present results.

6. EXPECTED CONTRIBUTIONS TO KNOWLEDGE

At end of this study, it is hoped that the following would have been achieved:

i. Add to the limited research available on security threats posed on social media to enterprise information systems in a developing country like Nigeria.
ii. Development of a lexicon based sentiment analysis model for assessing security threats posed on social media.

REFERENCES

1. Abuzaakouk, A. (2015). Sentiment Analysis for Security Sector Reform. In *NDIA HSD Human Systems Conference* (pp. 1–27).
2. Ali, A. (2017). Sentiment Analysis on Twitter Data using KNN and SVM, *8*(6), 19–25.
3. Amaratunga, D., Baldry, D., Sarshar, M., & Newton, R. (2013). Quantitative and qualitative research in the built environment : application of `` mixed ' ' research approach Dilanthi Amaratunga. *Emeraldinsight:Information Management & Computer Society*, *51*(1), 17–31. https://doi.org/10.1108/00438020210415488
4. Barclay, F. P., Pichandy, C., Venkat, A., & Sudhakaran, S. (2014). TWITTER SENTIMENTS : PATTERN RECOGNITION AND POLL PREDICTION. *Emeraldinsight:Information Management & Computer Society*, *11*, 141–167. https://doi.org/10.1108/S2050-206020160000011017
5. Bauwelinck, N., & Lefever, E. (2019). Measuring the Impact of Sentiment for Hate Speech Detection on Twitter, (c), 17–22.
6. Cambero, A. (2016). A Comparative Study of Twitter Sentiment Analysis Methods for Live Applications.
7. Martins, R., & Gomes, M. (2018). Hate speech classification in social media using emotional analysis, (April 2019). https://doi.org/10.1109/BRACIS.2018.00019
8. Subramaniyaswamy, V., Logesh, R., Abejith, M., Umasankar, S., & Umamakeswari, A. (2017). Sentiment Analysis of Tweets for Estimating Criticality and Security of Events. *Journal of Organisational and End User Computing*, *29*(4), 51–71. https://doi.org/10.4018/JOEUC.2017100103
9. Tang, Y., & Dalzell, N. (2019). Classifying Hate Speech Using a Two-Layer Model. *International Journal of Statistics and Public Policy*. https://doi.org/10.1080/2330443X.2019.1660285
10. Yuan, B. (2016). Sentiment analysis of twitter data, *2016*(March).