Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.9, No.8, 2018

A Framework for Privacy and Hate Speeches Control on Social Media in Mobile Computing

Souley Boukari[1*] Ismail Abdulkarim Adamu [2]

1. Department of Mathematical Sciences, Abubakar Tafawa Balewa University, Bauchi, Nigeria
2. Department of Computer Science Technology, Gombe State Polytechnic Bajoga, Gombe, Nigeria

* E-mail of the corresponding author: adamuismail200@yahoo.com

Abstract

The hype in information technology has led to the use of mobile devices to carry out several activities such as financial transactions, communications, learning and other related activities using the internet. Social networking platforms today are used to facilitate the activities carried out on mobile devices. The use of these social network platforms such as Facebook, Twitter, Whatsapp, Snapchat, Instagram among others on mobile devices is called mobile social network. A mobile social network makes it faster and easier for users to relate to one another and carry out other related activities at convenience. The major challenge with the use of social network applications on mobile devices is that, many illegal activities are carried out such as hate speeches and bigotry statement which has demeaning the interest of users. This research work proposes a technique for discovering and blocking hate speeches and bigotry users on social network platforms using text scanning and matching algorithm; in order to curtail the damages caused to national security and unity by such statement.

Keywords: Mobile social network, social networks

1. Introduction

The advent and use of internet in this twenty first century has made it easy for people in different parts of the world to interact and exchange meaningful ideas easily. The architecture of the internet is so flexible and fast that people can communicate sensitive and vital information around the globe within the blink of the eye (Tonge *et al*, 2013). To enjoy from the diverse benefits of the internet, social networking applications such as Facebook, twitter, Whatsapp and several others were developed and being used (Muse, 2014).

These social network applications have carouse a lot of addictions by both business men to promote their businesses and politicians to convey their political interest, communicate with citizens and individual who use them to meet new friends around the world (Jabee & Alam, 2016; Singh & Singh, 2017). These social networks platforms are today mostly used on mobile devices for easy accessibility and utilization, a new development called mobile social networks (Ometov *et al*, 2017,). Mobile social networks today makes it easier for people to access and enjoy the dividends of social networks for transactions, learning, communications and other related social activities at convenience and without delay (Mirzoev *et al*, 2014; Singh & Singh, 2017).

While the advantages of the internet and social networks cannot be over emphasized, some individuals have adopted the tool for social engineering thereby causing harm with it rather than good. Many cybercrime such as hacking, theft, cyber stalking, identity theft, malicious software, child soliciting and abuse are perpetrated on this platform. Other common cases in Africa, most specifically Nigeria is hate speeches and bigotry statement which has drawn government attention in recent times are executed on social media (Reddy & Reddy, 2014). These threats posed by malicious users have caused so much harm to different countries must especially Nigeria in terms of security, economy and the unity of the nation at large. The need to scan, match and filter out hate speeches and bigotry statement using appropriate technique on social networks platform become imperative in order to combat it rampage. The remaining part of the work is organized as follows section 2 related work, section 3 methodology, section 4 expected result and section 5 conclusion.

2. Literature Review

Social media, most especially Facebook, is very useful today in creating a platform to share ideas and meet new friends around the globe. However, the medium has been hijacked by different individuals, who hurl abusive word to friends or other users on the platform. In view of that, Ekwueme and Anthony (2017), conducted a review to identity the excess of abuse and fowl languages on social media. In their research, they discovered that, the abuse of social media is in the increesaes and very little or no effort are made to check mate them. They finally urge the government to riseup against the abuse of the social media platform and also train personal that

can detect and prevent the excesses. Also, David and Fernandez (2016) considered ways that overt and covert descrimation practices that circulates on facebook despite of its offical policy that prohibit foul word and hates speeches. Their research analyzed different data extracted from offical facebook pages of seven extreme right political parties in spain between 2009 to 2013. They discovered in their nalysis that, spain extreme right politcal parties primarily implicate discrimation which is then seized by their followers, who vehemently overt hate speeches in the comment space on the social media. It was concluded in their research that, except new policy are introduce to curtail the way hate speeces and descrimination are seized and exhibited on Facebook, prevalence of hate speeches will continue to be in circultion on users platform. Moreover, Alkali *et al*, (2017), examine the phenomenon of hate speeches and foul languages on social media platforrms in Nigeria in order to access their moral and legal consequecies in the society and to journalism practice. Their findings identfied that, hate speeches and foul languages on social media have moral and legal consequencies on the society and practice of journaism. They concluded base on their discovery that, eventhough there are legal laws that can be use to curb hate speeches and foul language on social media platform in Nigeria, still, they urge Nigerian government and NGOs to consolidate effort in installing a monitoring team that would identify and remove hate speeches content on social media platform in Nigeria. Furthermore, to curb outspread of hate speeches on social media, Santos *et al*, (2018), introduced and approached to tackle the problem of offensive anguages in online social media. Their proposed techniques, uses an unsupervised text style transfer to translate offensive languages or sentencies in to a non offensive one. Additionally, Kumar *et al*., (2018), carried out a shared task on aggression identification in order to develop a classifier that would descrimate between overtly aggressive, covertly agrressive and non aggressive text on social media platform. To achieve this, participant were trained in Englsih and Hindi with 1,500 dataset of annotated Facebook post and comment. Their result depict how tedious and difficult it was to identified hate speeches on socil media platform. However, research are ongoing in a quest to have a mechanish that will detect and fetch out hate speeches and abusive users on social media platform in order to curtail the excesses of hate speeches on social media latform.

3. Methodology

This research work proposed to develop and implement a framework for privacy and hate speeches control mechanisms on social networks platforms using text matching and scanning algorithm. The comments added by each user or posted on his wall and other users wall will be first of all scan through using the text scanning algorithm and matched with the available hates and bigotry statements character in order to detect hates speeches and bigotry statement, which is reported to the admin for necessary action. The proposed system will be developed using Java programming tool on netbeans platform to test the algorithm on a real chat system.

3.1 Working principle of the proposed system

The working principle of the proposed system is described as follows:

1. Start
2. Add comments/post
3. Scan comments
4. Filter and match bigotry statement and hates speeches
5. If bigotry statement and hates speeches are discovered **THEN**
6. Obtain user information, report user to admin and bloc user account **ELSE**
7. Allow user to continue operation
8. Stop

3.2 Architecture of the proposed system

The architecture of the proposed system is shown in figure1 below:

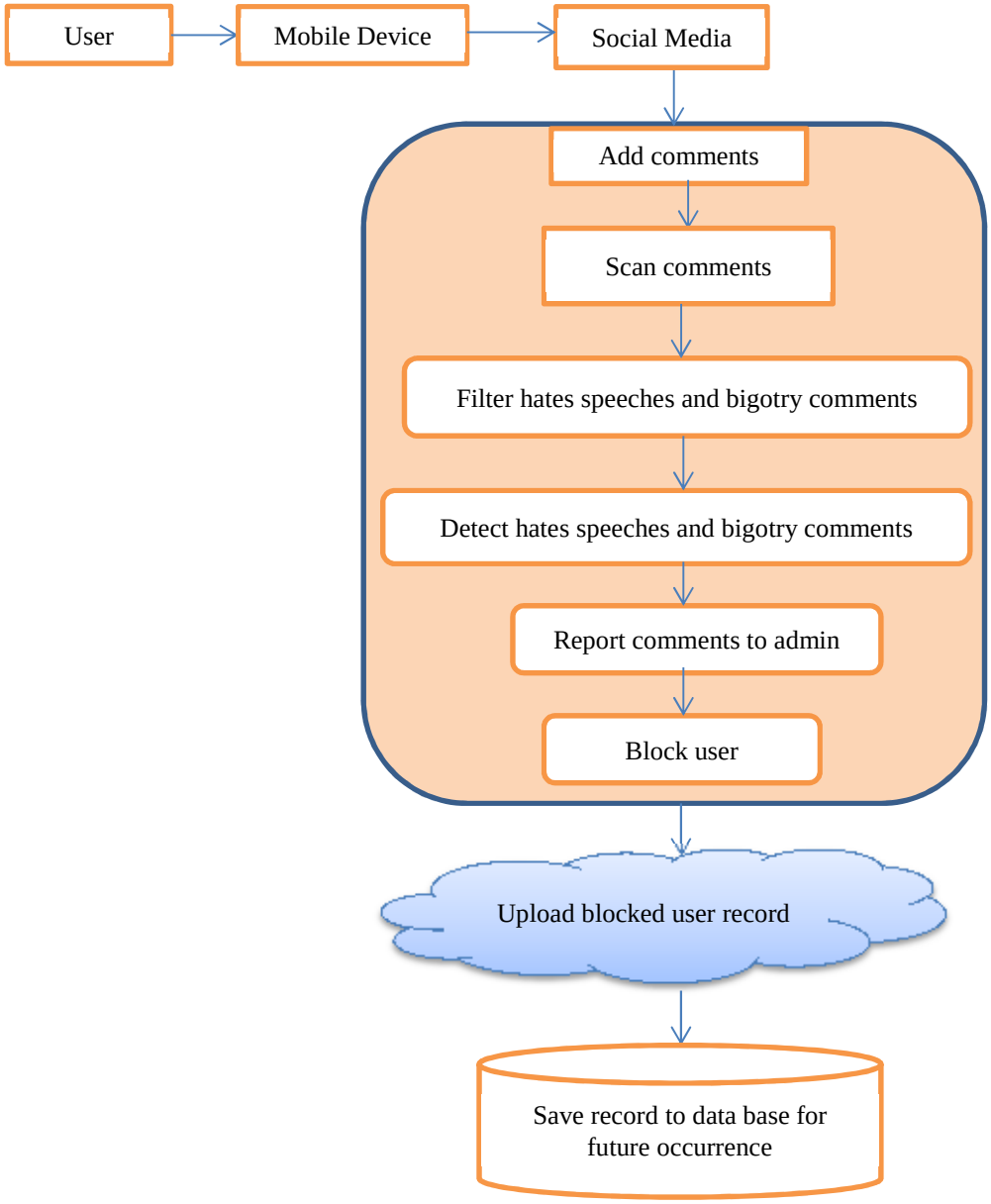

Figure1: Architecture of the proposed System

4. Expected Result

This research work intends to come up with an improved privacy and hate speeches control mechanism on social networks application. This will help in identifying the source, tracing individuals who resort in using the social media platform in perpetrating hate speeches and bigotry statement. The intended application will be achieved through the implementation of text scanning and clustering algorithms.

5. Conclusion

The use of the internet on mobile devices and the addiction of social networks platforms on such devices by politicians, business organizations, institutions and different individuals to carry out their various activities are on the increase today. However, while many are utilizing the benefits of the internet and social networks for profitable activities, others have resort to abuse the opportunity by posing security threats and unleashing bigotry activities which have raised a lot of eyebrow by users, security agencies and government at large. In view of the above challenges this scientific research work aim to implement a framework for privacy and hates speeches control mechanisms on social networks platforms using text scanning and clustering algorithm to curb the menace of these illegal activities by the perpetrators. This research work will reduces the high rate and spread of hates speeches and bigotry statement on social media platforms which has maim users in recent times when

implemented. In future, we intend to implement and test run the propose algorithm on a real chat system to validate the algorithm and monitor it operation.

References

Alkali, T. T., Faga, P. H., & Mbursa, J. (2017). Audience Perception of Hate Speeches and Foul Languagge in the Social Media in Nigeria: Implications for Morality and Law. *International Scientific Journal*, 161-178.

Chen, Y., Zhou, Y., Zhu, S., & Xu, H. (2012). chen_etal_socialCom. Retrieved October 19, 2018, from www.computer.org: https//faculty.ist.psu.edu

David, A. B., & Fernandez, A. M. (2016). Hate Speech and Covert Discrimination on Social Media: Monitoring the Facebook Pages of Extreme-Right Political Parties in Spain. *International Journal of Communication*, 1167-1193.

Ekwueme, A., & Ugwuta, C. (2017). Appraisal of the Excess and Abuses of the Social Media in Contemporary Society: The way Forward. *Global Journal of Huma-Social Science: A Arts & Humanites -Psychology*, 17(4), 1-8.

Frank, I. D., & Odunayo, E. (2013). Approach to Cyber Security Issues In Nigeria: Challenges and Solution. *IJCRSEE*.

Internet Live Stats. (2016, July 1). Nigeria Internet Users. Retrieved May 8, 2017, from internet live stats: internetlivestats.com/internet-users/nigeria/

Jabee, R., & Alam, M. A. (2016). Issues and Challenges ofCyber Security for Social Networking Sites (Facebook). *International Journal of Computer Appications*, 144(3), 36-40.

Kumar, R., Ojha, A. K., Malmasi, S., & Zampieri, M. (2018). Brnchmarking Aggression Identification in Social Media. Procceding of the First Workshop on Troling, Aggression and Cyberbullying, (pp. 1-11). Santa Fe.

Mirzoev, T., Brannon, M., Lasker, S., & Miller, M. (2014). Mobile Application Threats and Security. *World of Computer Science and Information Technology Journal*, 4(5), 57-61.

Muse, S. A. (2014). Social Media and public Participaton in Nigeria: Challenges and posibilities. *Advancesin Social Sciences Research Journal*., 1(7), 43-48.

Ometov, A., Levina, A., & Borisenko, P. M. (2017). Mobile Social Netwrking Under Side-Channel Attacks Practical Security Challenges. IEEE, 5, 2591-2601.

Reddy, G. N., & Reddy, G. J. (2014, January 30). A Study of Cyber security challenges and itsEmerging Trends on atest Technologies. arXiv preprint arXiv, pp. 1-6.

Santos, C. D., Melnk, I., & Padhi, I. (2018). Fighting Offensive Language on Social Media with Unsupervised Text Style Transfer. Proceedings of the 56th Annual meeting of the Association for Computational Linquistics (pp. 189-194). Melbourne: Association for Computer Linguistic.

Singh, A., & Singh, A. (2017). Review of Cyber Threats in Socia Networking Websites. *International Journal of Advanced Research in Computer Science*, 8(5), 2695-2699.

Tonge, M. A., Kasture, S. S., & Chaudhari, R. S. (2013). Cyber Security: Challenges for Society-iterature review. *Journal of Computer Engineering*, 12(2), 67-75.