A Graph based Cloned Profile Detection in Online Social Networks

Rahul Nair[*], Bhuvaneswari Anbalagan[**]

[*,**] School of Computer Science Engineering,
[**] Vellore Institute of Technology, VIT, Tamil Nadu Chennai

DOI: 10.29322/IJSRP.11.06.2021.p11484
http://dx.doi.org/10.29322/IJSRP.11.06.2021.p11484

Abstract- Social media is one of the most important part of an individual's life. A person spends time on social media for recreation or work related work. While these social media helps the person increase their network by either connecting their friends or colleagues. Social networks also help in connecting people with similar interests. One can post about their opinions about something and people can comment on the post and react to the post. People can also post photos and videos about their interests. With all the perks offered by the social networks, certain security issues arise due to the ease of availability of information about a person, identity of a user is cloned and used for malicious intents. The cloned profile could be used to actually harm the actual person's reputation. If the person is a popular person, the posts about a certain topic could have a significant response. For example, a politician posting about a certain topic is a adverse way may trigger the citizens of the country and there could be catastrophic consequences. Hence it is really important to detect such accounts on the social networks so as to prevent such catastrophic consequences and also keep the social network a healthy platform for people to interact and network. This research paper aims to detect such cloned profiles in Online social networks considering the accounts being graphs and the attributes being the nodes of the main node that is the account name and username.

Index Terms- Cloned Profile, Graphs, Online Social Networks, Reputation, Social Media

I. INTRODUCTION

Online Social Networks, also abbreviated as OSNs are social networks which consist a set of individuals, which are known as the nodes a of the social networks and the relationships between the nodes that is, the ties between individuals and individuals, over a common platform on the Internet. The relationships between entities or nodes is primarily based on similar interests be it personal or career, matching backgrounds or if the entities know each other in real life. Social Networks have different features and also differ in format. The can be operated on Desktop computers or via an application on the phone. Some of the common features of Online Social networks are Posting about a topic in which the individual's or the organization's interest lies. the content shared by another entity if found interesting, informative or relatable. Communication between entities via text messaging, voice calls or video calls.
Different Online Social Networks have different focuses, that is, the primary objectives of the OSN. One such categorization of objectives of different Online Social Networks is as follows Socializing Based: Online Social Networks like Facebook are primarily aimed to connect people for socializing that is, the services of Facebook are used to connect with existing friends of a person. Non Socializing Based: Online Social Networks are also used for non-social communication between entities that is, the purpose of interaction is purely business. An example of such an OSN is LinkedIn. Social Navigation Based: The services of such OSNs are mainly used for heling other users to find specific data or resources.

Facebook reported almost 1.85 billion daily active users during the fourth quarter of 2020[1]. The total number of users on Instagram was estimated to be 1.16 billion active users.[2] LinkedIn stated that it had 740 million users in more than 200 countries and territories all around the world in 2021.[3] So here we can see that, a lot of people and organizations use Online Social Networks for reasons like recreation, business or just merely connecting to their old friends who they cannot meet in real life. Online Social Networks have now become an integral part of communications and interactions in today's world. An entity uses social media to connect to friends, connecting to long lost friends, setting up communities based on businesses or simply for recreation. To become a part of an Online Social Network like Facebook, LinkedIn, Twitter, the user has to enter personal information which makes it easier to connect to people on these platforms. But with all the perks which come with Online Social Networks, there's problems that arise with sharing all that information. People may use the information to create profiles that imitate the original users and can dupe other people into believing them as original users and then use the connections for malicious purposes. Malicious intenders can use the cloned profile to spread hate content, opinionate posts which could lead to controversies and which in result could harm the reputation of the person of whose the profile is cloned. Profiles that are intended for malicious purposes usually have the same name as the original profile so as to get as much attention as possible as the original profile might in order to maximize the damage done. Also, the malicious intender could use the fame and reach of the profile in order to manipulate financial markets so as to get profit while on the

cost of a lot of people losing their money. For example, a person trying to clone Elon Musk(CEO, Tesla Motors) could use the fame Elon Musk has to manipulate cryptocurrency markets and could lead to bull or bear runs. This could result in a lot of people losing a lot of money. Another example would be someone impersonating a political leader like Narendra Modi. A cloned profile of Narendra Modi(Prime Minister of India) could easily instigate a group of people by posting triggering political opinions on sensitive political or diplomatic matters. A person cloning to be Marquez Brownlee(a popular YouTuber who posts videos about reviews of devices and also other technology related videos) could false reviews of a device, that is, if he posts good reviews of a device that is underperforming, he will increase the sales by duping the people while on the other hand, if the device is performing good and the reviews about the latter are not good, it may decrease the sales of the product which in turn would result in a loss for the company. These examples explain how dangerous profile cloning could be for a person's reputation and also for the smooth functioning of different societies that is, economic, social and political societies. Profile cloning has become really common in Online Social Networks due to the ease in getting the information about the user. Even after the privacy measures taken by Online Social Network platforms like Facebook, Instagram, Twitter, still the information of a person for example his name, DOB, profile photo can be easily accessed and hence can be used in order to create a profile for purposes of their own.

Apart from profile cloning, there are other threats in Online Social Networks. Some of the threats are mentioned below:Cla ssic Threats: Malware: Malware is a commonly used term for malicious software that is designed to harm or exploit any device or network. Malware is generally used to get data from the affected device and use that data for malicious purposes or for monetary gains. The links could be provided on social media posts and users who click on those links would have the malware downloaded onto their machine or network and then, the malware will be used for purposes mentioned above. Phishing: Any malicious intender could impersonate reputed organizations or an employee of such organizations in order to get control of sensitive information about a person or another organization which also could be used for malicious purposes or could be used as leverage in order to get money from the affected party/parties. Cross Site Scripting(XSS): Cross Site Scripting is a code injection attack that happens on the client side. The main aim of the attacker is to execute malicious scripts on the web browser of the victim by including the malicious code in a legitimate looking web page or web application. Cross Site Scripting along with social engineering could enable criminals to pull of attacks like keylogging, phishing, planting trojans and data theft.

Clickjacking is an attack a link of an element on a webpage which is usually disguised as another element which causes the victim to download unwanted software or maybe malware, visit malicious web pages or even provide sensitive information which could in turn lead to monetary loss. Other than that, the malware downloaded can be used for purposes mentioned above. Social Bots: Social bots are accounts that are not a real user. These accounts are used to spread false news or rumors, spread opinions about a sensitive matter, spam a person's profile or to defame a person online. Since these are not real users, there are a lot of programmed bots on social media that are used to accomplish the above said tasks and a bunch of bots can be controlled by a single person. Social Bots: Social bots are accounts that are not a real user. These accounts are used to spread false news or rumours, spread opinions about a sensitive matter, spam a person's profile or to defame a person online. Since these are not real users, there are a lot of programmed bots on social media that are used to accomplish the above said tasks and a bunch of bots can be controlled by a single person. Personal Threats: Bullying someone over the Internet is a common nowadays. Malicious intenders bully people by fat shaming them, giving them death threats and/or threats against their families. This bullying is done wither for the recreation of the bully or to make the victim do something which the victim is unwilling to do.

II. RELATED WORK

Mateen et al(2017)[4] divided the features of a user on a social network into three different types; User Based Features: These features included the relationships user has with other users on the social network, the number of followers the user has, the number of people the user follows, the age of the account that is, the time passed after the creation of the account and the ratio of number of accounts that the user follows to the number of accounts that follow the user. Content Based features: These features include the content posted by the user that is, for example, the tweets. The content also includes the hashtags in each post, the URLs the user posts and the frequency by which the user posts a tweet. Graph based features: These features include certain parameters like in/out degree(that is the ratio of the number of incoming links to a particular node to the number of outgoing links from the same node in a graph), Betweenness Centrality(a measure of centrality which is based on shortest path. For every pair of vertices in a connected graph, there exists at least one shortest path in between the vertices in such a way that the sum of weights of the edges or the path that is taken from one vertex to another is minimized. The betweenness centrality for each vertex is the number of the mentioned shortest path that pass through the vertex.

Mohammadrezaei et al(2018)[5] proposed a way to detect fake accounts by computing similarities between the accounts concerned. The proposed way was to initially create a sparse adjacency matrix of the social graphs, then a similarity matrix was computed for measures like Common friends, Total friends, Jaccard similarity, Cosine similarity, L1 Norm similarity. Zaré et al(2018)[6] proposed a methodology that uses the user's followers count as a metric and uses a clustering algorithm known as K Nearest Neighbors paired along with the PageRank algorithm to detect cloned profiles. The methodology says to initially measure the user's follower count and then KNN clustering is user to group the users with other profiles having similar attributes. Then based on the PageRank algorithm, real famous users, who have many followers but follow a small number of users gain a high rank from these relationships. Mohammadrezaei et al(2019)[7] proposed a similar methodology as in (b) but with a few set of changes for the detection of fake profiles. Initially as in (b), a sparse adjacency matrix is created for the social graphs. Now distinctive attributes are selected which would help detect fake accounts. After this step, several matrices that represent similarity between the nodes were

shown. The dataset used in the paper was biased as there were more legitimate users than fake users so the dataset has to be balanced. To balance the data, SMOTE algorithm was used. Van Der Walt et al(2017)[8] did a study on how the existing methods to detect bots on social media would work on detecting fake human profiles in these social networks. Principal Component Analysis was then used to select the top 10 columns with the highest variance from each of the matrices and the data tags were applied that is, the data was labelled. After that, supervised learning algorithms like Linear Support Vector Machine, Medium Gaussian Support Vector Machine and Logistic Regression was used for classification[9], [10]. The balancing of the dataset gave a better performance to the classifier[11].

On the other hand, cloned profiles have low ranks [12] because, despite having inbound links, they have not been followed by other famous users who have higher ranks [13]. Furthermore, they are prone to have a lot of outbound links too[14]. The cloned profiles, therefore, gain lower ranks and could be easily identified based on their rank[15]. The advantages of the above methodology is that the approach can leverage the power of Big data analytics[16], semantic[17], to work on large datasets and hence [18], it is efficient[19]. Also, the methodology proposed does not require any human intervention. Now, for number of mutual friends, total number of friends, Common neighbor, Common neighbor graph edges, Jaccard similarity, Cosine similarity and other measures, the similarity matrix is computed[20]. Then using Principal Component Analysis, the initial space was reduced to columns having informative values and then using the elbow method, a set of highly informative values is selected. Now, Stochastic Support Vector Machine(SVMSch) and Neural Networks [21] are used to train the machine and then the thresholds are computed. The legitimacy of the accounts is determined by entering the new data with respect to the threshold values. The problem with the approach is the time taken to execute the entire system that is, time complexity for extracting each similarity matrix is O(n2). PCA is applied to each similarity matrix and the time complexity of PCA is equal to O(min(n3, p3)) where n is the number of samples and p is the number of features. Since for this scenario, p is equal to n, the time complexity becomes O(n3). Next, in the learning step, the time complexity for OSVM(One Class Support Vector Machine) is O(n2) so the net time complexity turns out to be O(n3). The machine learning model that was previously used for the detection of bots on social media to detect deceptive human accounts was taken into consideration [22]. Then the bot accounts were cleared out from the dataset and 15000 fictitious accounts were created and added to the dataset. Now the classifier was fed engineered features that is, features that came out as a result of combining the existing features and the results were computed and evaluated. The conclusion of the work was that the existing techniques that were used to detection of bots were not good for classifying deceptive human profiles in these social networks.

Now it is the time to articulate the research work with ideas gathered in above steps by adopting any of below suitable approaches: The cloned profile could harm the reputation of the original user and that could lead to problems for the victim. For instance, if the cloned profile is spreading news that is unreliable or fake, the people will stop relying on the legitimate news provided by the actual user. The cloned profile could manipulate financial markets if the profile that is cloned is of a well-known person in the financial industry. It could cause panic selling or buying which would benefit the malicious intender but could cause a lot of people to lose their money. The cloned profile could instigate citizens by sharing offensive or triggering posts from a cloned profile of a well-known person in politics which could lead to unrest or even defaming the victim. Defaming the victim could cost the vote of the citizens in the next upcoming elections. Cloned profiles of well-known product reviewers or product enthusiasts could give out fake reviews which could lead to the increase in sales of a bad product leading to people buying products that are not at par with the industry standards leaving the customers not satisfied while also the profile could bring down the sales of a good product which stops a good product to as many customers as possible.

III. PROPOSED SYSTEM DESIGN

The idea of the system is to visualize the account of a person as a graph and then based on the graph attributes that are the profile attributes of an account to create a behavioral and a attribute based profile which would represent a person and the things the person likes to talk about. When a malicious intender tries to clone the profile of a person for malicious purposes, the profile is checked against the general characteristics of the original profile and then it is flagged as cloned based on the attributes compared and the similarity of the tweets. The system initially takes in the user name of the 2 profiles, That is, the legitimate account and the account that is suspicious of being the clone of the original account. Then the user attributes of both the accounts are scraped from the social media website using a tool known as Twint. Twint is a tool which is used in this research for retrieving user data and tweets from Twitter. Twitter is the social media website we are using for carrying out the research. Although, the system that is proposed in the report is completely capable of detecting cloned profiles in other social networks. For example, Facebook, Instagram and LinkedIn. The reason for using Twitter is the ease of collecting data about a user from Twitter. As long as the data required by the system is given to it, this system can be applied to any social network. After the user attributes are retrieved, certain attributes are taken in to consideration. That is, the Screen name of the profile, the number of accounts following the concerned account, the number of accounts the concerned account follows and the location of the profile at the time of creation. Then Twint is again used to scrape the tweets of the 2 profiles that are considered. The tweets are then stored in a data frame after which the tweets are pre-processed. Pre-processing the tweets involves the removal of punctuation, the removal of special symbols and removing anything that are not relevant to the tweets. Pre-processing also includes converting all the words in the text to lower case.

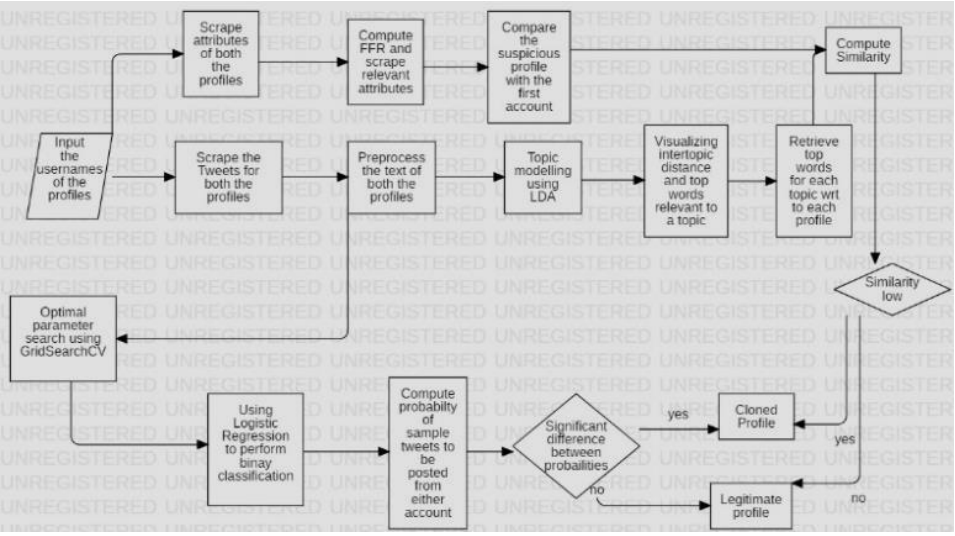

Figure 1. Proposed System Design

After the pre-processing, for the nouns and adjectives are used to define common topics for the tweets by a particular person. Initially, the original text was taken into account, the results were not satisfactory. Then the words were further refined and the model was based on nouns. It gave better results but then again, it was not up to the mark. The topics that were the result of the model were not differentiable. Then nouns and adjectives were used to which the results were really good. For topic modelling, Latent Dirichlet Allocation was used which is a topic modelling technique. The topics were modelled and then the top words in each topic were stored in a dictionary as well as visualized with relevant terms per topic. This step was done for both the accounts. The top 20 words for each topic were taken and stored in the dictionary and then each topic of 1 account was compared to all the topics of the other account to compute the similarity between the topics. If the similarity was too low, the suspicious account can be flagged as a potential cloned account with malicious intent. Similarity was computed using the cosine similarity measure

$$\text{Similarity}(A,B) = A*B/(||A||*||B||) \quad (1)$$

Post this, the tweets are combined into a data frame. Logistic Regression was used to classify the accounts on a probability for the two labels that is the name of the first account and the name of the second account. The probability is measured for each account being the author of a tweet. If the accounts are operating by the same person, the account is legitimate and hence, the tweet probability will be equally distributed while if the account is cloned and used by another person for malicious purposes, the probability is highly different. Hence all these metrics can be used to flag the account as legitimate or cloned.
Logistic regression follows the below formula:

$$y = e(c+m*x)/(1+e((c+m*x)) \quad (2)$$

where e is again, the Euler's number equivalent to 2.71828(upto 5 decimal places) or again, also known as the base of natural logarithms, c is the bias or the intercept and m is the coefficient for the input value x.

IV. CONCLUSION

The proposed system, visualizes social media account as a graph and the attributes of the accounts as nodes that are connected to the main Node, that is, the account. When a malicious intender tries to clone the profile and tries to hinder the reputation of the original account, the attributes of the account differ and the way the malicious intender posts is different from the posts of the actual person. When a profile is cloned, the Screen name of the original account is usually copied so that people on social media are fooled into following the cloned profile of the person. The system leverages the difference in the above mentioned attributes to determine if the profile is a legitimate profile or a cloned profile for malicious intent. The results from comparing the 2 profiles in each of the test case shows that if the account is a cloned account of a person, the attributes like the Following-Follower Ratio ratio, the location, the average number of retweets are highly different. Also, the similarity measure between the topics of the 2 accounts is really less too. Moreover, when the tweets posted by the 2 accounts are checked, there is a vast difference in probability of the 2 tweets being posted by either of the user. The FFR (Following-Follower Ratio is calculated by: FFR is the N1 Number of accounts that the profile follows to N2 Number of accounts that follow the concerned profile.

$$FFR = \frac{N1}{N2} \quad (3)$$

The average retweet count of a user is calculated by the following formula. Hence the Average Retweet Count is the ratio of Number of retweets (n) per tweet posted by the user to Total number of tweets extracted of the profile(N).

$$AVG(RT) = \frac{n}{N} \quad (4)$$

TABLE I.
Confusion matrix for the logistic regression model of JoeBiden and biden4pres

	Positive	Negative
Positive	493	32
Negative	80	471

TABLE II.
Confusion matrix for the logistic regression model of KamalaHarris and KamalaHarrisVP3

	Positive	Negative
Positive	534	28
Negative	88	536

The logistic regression model was tested with various account pairs on Twitter having the same screen name. The logistic regression model gave a 90.21% accuracy, 85.85% precision, 95% recall and a F1 score of 0.90 for the account of Kamala Harris which had the username KamalaHarris and an account which had the name Kamala Harris but with the username KamalaHarrisVP3. Similarly, for another example, the model was checked for the account pair of US President Joe Biden with the username as Joe Biden and another profile with the name Joe Biden but with the username biden4pres.

TABLE III.
Scores for the logistic regression model

	JoeBiden and biden4pres	KamalaHarris and KamalaHarrisVP3
Accuracy	0.8959	0.9022
Precision	0.8600	0.8585
Recall	0.9390	0.9502
F1	0.8979	0.9020

The model output with an accuracy of 89.59%, precision of 86%, recall of 93.9% and a f1 score of 0.897. To check for actually similar accounts, simulation was done by sampling was done for the tweets of the 2 accounts and then checked against the same measures. The similarity between tweets came out to be high when compared to one another and the probabilities of the tweets being of either class was nearly equal as well. The system for checked for accounts with the same Screen Name and the system gave desired satisfactory results.

V. CONCLUSION AND FUTURE WORK

A conclusion section is not required. Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions. The proposed system, visualizes social media account as a graph and the attributes of the accounts as nodes that are connected to the main Node, that is, the account. When a malicious intender tries to clone the profile and tries to hinder the reputation of the original account, the attributes of the account differ and the way the malicious intender posts is different from the posts of the actual person. The project that is specified currently only uses English language to model topics based on the tweets and check probability. The system faced a challenge when other languages were used. In the future, the system could be added the support of multiple languages in order to check cloned profiles in different regions as well. Another challenge that was faced was the speech in other languages written in English, that is, for example, Hindi words in English which makes it difficult for modelling topics or machine learning algorithms to pick up. Further work on the system could add features that would possibly translate the words to a language that the machine understands. Another future prospect of the system would be to add image recognition features for comparing the images posted by both the accounts to detect the dis similarity in case of cloned accounts of similarity in case of legitimate accounts..

REFERENCES

[1] https://www.statista.com/statistics/346167/facebook-global-dau/

[2] https://www.searchenginejournal.com/instagram-facts/314439

[3] https://www.oberlo.in/blog/linkedin-statistics

[4] M. Mateen, M. A. Iqbal, M. Aleem and M. A. Islam, "A hybrid approach for spam detection for Twitter," 2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST), 2017, pp. 466-471, doi: 10.1109/IBCAST.2017.7868095.

[5] Mohammadreza Mohammadrezaei, Mohammad Ebrahim Shiri, Amir Masoud Rahmani, "Identifying Fake Accounts on Social Networks Based on Graph Analysis and Classification Algorithms", *Security and Communication Networks*, vol. 2018, Article ID 5923156, 8 pages, 2018.

[6] Zaré, Maryam et al. "Automatic ICA detection in online social networks with PageRank." *Peer-to-Peer Networking and Applications* (2020): 1-15.

[7] Mohammad Reza Mohammadrezaei; Mohammad Ebrahim Shiri; Amir Masoud Rahmani. "Detection of Fake Accounts in Social Networks Based on One Class Classification". The ISC International Journal of Information Security, 11, 2, 2019, 173-183.

[8] E. Van Der Walt and J. Eloff, "Using Machine Learning to Detect Fake Identities: Bots vs Humans," in *IEEE Access*, vol. 6, pp. 6540-6549, 2018, doi: 10.1109/ACCESS.2018.2796018.

[9] Devakunchari Ramalingam, Valliyammai Chinnaiah, Fake profile detection techniques in large-scale online social networks: A comprehensive review, Computers & Electrical Engineering, Volume 65, 2018, Pages 165-177, ISSN 0045-7906

[10] Nguyen H.A., Nguyen T.T., Pham N.H., Al-Kofahi J.M., Nguyen T.N. (2009) Accurate and Efficient Structural Characteristic Feature Extraction for Clone Detection. In: Chechik M., Wirsing M. (eds) Fundamental Approaches to Software Engineering. FASE 2009. *Lecture Notes in Computer Science*, vol 5503. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-00593-0_31

[11] P, Sowmya and Chatterjee, Madhumita, Detection of Fake and Cloned Profiles in Online Social Networks (March 9, 2019). Proceedings 2019: Conference on Technologies for Future Cities (CTFC).

[12] P. Bródka, M. Sobas and H. Johnson, "Profile Cloning Detection in Social Networks," *2014 European Network Intelligence Conference*, 2014, pp. 63-68, doi: 10.1109/ENIC.2014.21.

[13] Bhuvaneswari, A., and C. Valliyammai. "Information entropy based event detection during disaster in cyber-social networks." *Journal of Intelligent & Fuzzy Systems*, IOS Press 36, no. 5 (2019).

[14] Valliyammai. C., Bhuvaneswari A., "Semantics-based sensitive topic diffusion detection framework towards privacy aware Online Social Networks", *Cluster Computing, Springer* (2018) pp. 1-16

[15] Bhuvaneswari, A., J. Timothy Jones Thomas, and P. Kesavan. "Embedded Bi-directional GRU and LSTM Learning Models to Predict Disasterson Twitter Data." *Procedia Computer Science, Elsevier* 165 (2019)

[16] Bhuvaneswari. A., Karthikeyan, M., Lakshminarayanan, T., "Improving diversity in video recommender systems and the discovery of long tail", *Journal of Theoretical and Applied Information Technology*, Vol.37, Issue.2, pp. 224-233, 2012.

[17] Bhuvaneswari, A., K. Aishwarya, S. Bhuvaneshwari, C. Sai Chandni, and P. Sundara Akilesh. "Detecting New Events from Microblogs Using Convolutional Neural Networks." T. Sengodan et al. (eds.),*In Advances in Electrical and Computer Technologies, Lecture Notes in Electrical Engineering, pp. 1-9. Springer, Singapore*, 2020. 672.

[18] Bhuvaneswari, A & Valliyammai, C, 'Semantic-based sensitive topic dissemination control mechanism for safe social networking', Rajsingh, E.B, et al. (eds.), *Advances in Big Data and Cloud Computing, Advances in Intelligent Systems and Computing*, Volume 645, Chapter No. 17, pp. 197-207. Springer Nature, Singapore, 2018.

[19] Bhuvaneswari, A & Valliyammai, C, 'Social IoT enabled emergency event detection framework using geo tagged microblogs and crowdsourced photos', Abraham, A, et al. (eds.), Emerging Technologies in Data Mining and Information Security, *Advances in Intelligent Systems and Computing*, Volume 813, Chapter No. 13, pp. 151-162. Springer Nature, Singapore, 2018. DOI: 10.1007/978-981-13-1501-5_13.

[20] Bhuvaneswari A., Valliyammai. C., "Identifying event bursts using log-normal distribution of tweet arrival rate in Twitter stream", IEEE 10th International Conference on Advanced Computing (ICoAC), MIT Campus, Anna University, Chennai. 13-15 December 2018.

[21] Bhuvaneswari A., Valliyammai C, Devakunchari R, "Feature Constrained Parallel Data Processing Approach for Spatiotemporal Event Detection", IEEE 9th International Conference on Advanced Computing (ICoAC), MIT Campus, Anna University, Chennai. 14-16 December 2017.

[22] Bhuvaneswari A., Valliyammai C, "#ChennaiFloods: Leveraging Human and Machine Learning for Crisis Mapping during Disasters Using Social Media." IEEE 23rd International Conference on High Performance Computing Workshops (HiPCW), Hotel Novatel, Hyderabad, IEEE, 19-22 December 2016.

AUTHORS

First Author –Rahul Nair, B.Tech, Vellore Institute of Technology, Chennai, rahulsnr98@gmail.com.

Second Author –Dr. Bhuvaneswari Anbalagan. Currently working as Assistant Professor (Senior Grade) in School of Computer Science Engineering, VIT University, Chennai, India. bhuvana.cse14@gmail.com

Correspondence Author – Dr. Bhuvaneswari Anbalagan. Assistant Professor (Senior Grade) in School of Computer Science Engineering, VIT University, Chennai, India, bhuvana.cse14@gmail.com