A Credibility Analysis System for Assessing Information on Twitter

Abrar Ahmad Mir, Er. Kirti Joshi, Dr. Ashish Oberoi
Department of Computer Science Engineering , RIMT University
Opposite Floating Restaurant, Sirhind Side, Mandi Gobindgarh-147301,
Punjab (INDIA)

***Abstract*:- Online social communities, for example, Twitter, have become exceptionally famous in the modern era, as the quantities of clients who are utilizing them on everyday increasing continuously. Data spread through these stages is their most alluring element, as it is known to be expedient and financially savvy. The way that clients are permitted to convey what needs be with next to zero control is likewise another exceptionally alluring part of these stages. As clients are managed the flexibility to distribute content with no supervision, the issue of data believability on informal organizations has likewise ascended as of late. The clients of these stages can spread in-arrangement maliciously for reasons that may not be perfect with the benefit of society. Clients are getting to be attentive that bits of gossip that are spread through online informal organizations can have impeding impacts due to there is no effective system exist for classify these suspicious tweets. In this paper we have proposed an approach for the classification of suspicious or threat related tweets utilizing k-Means clustering and k-NN classification algorithm.**

Keywords: Social Threat, twitter4j API, classification, k-Means and K-NN.

INTRODUCTION

In recent years, the different social networks sites gained popularity among peoples for the ability to connect and provide communication between peoples. These web based application also provide a way to share individual's experience, feelings to other peoples. Nowadays these social resources are used on-line interactions and content sharing like opinions, feelings, and sentiment expressions and so on. Twitter is most widely used sharing tool these days where people around all over the world can share there feeling or opinion through tweets and its use is increasing day by day. Many commercial organisations and government departments follow this social site to extracts useful data or detect suspicious activities utilizing different data mining tools. Data mining tools helps in extracting needed, useful and accurate information from these social resources by processing and analysing the available information. These data mining tools utilize information retrieval and machine learning and used to mine patterns, provide useful data and social site are very powerful places to apply data mining tools.

Data Mining is characterized as separating the data from the tremendous arrangement of information. As it were we can state that information mining will be mining the knowledge from information. This data can be utilized for any of the accompanying applications:

- Market Analysis
- Fraud Detection
- Security application
- Customer Retention
- Production Control
- Science Exploration

SOCIAL media sites such as Twitter provide great venues for criminal minded to share joy and struggle, vent emotion and stress, and seek social support. On various social media sites, criminal minded discuss and share their everyday encounters in an informal and casual manner. Criminal minded' digital footprints provide vast amount of implicit knowledge and a whole new perspective for educational researchers and practitioners to understand criminal minded' experiences outside the controlled classroom environment. This understanding can inform institutional decision-making on interventions for at-risk criminal minded, improvement of education quality, and thus enhance student recruitment, retention, and success. The abundance of social media data provides opportunities to understand criminal minded' experiences, but also raises methodological difficulties in making sense of social media data for educational purposes. Just imagine the sheer data volumes, the diversity of Internet slangs, the unpredictability of locations, and timing of criminal minded posting on the web, as well as the complexity of criminal minded' experiences. Pure manual analysis cannot deal with the ever growing scale of data, while pure automatic algorithms usually cannot capture in-depth meaning within the data.

In this paper we proposed a k-Means clustering and KNN classification based Twitter Tweets classification into five categories as follows Murder, Kidnap, Corruption Charges, Robbery and Theft and Match Fixing.

OBJECTIVES

1. To configure, conceptualize, and build up a Twitter Sentiming Analysis Tool for powerful Sentiment Analysis.

2. To fill in as a powerful Market Research Analysis device for any Company/Institution or Government which wishes to decide the input or client sentiment or the suspicious behavior of specific.

3.The guideline target of this proposed framework is to group the tweets from twitter between a hazard tweet and

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

not a threat tweet. This is done by utilizing a game plan of watchwords having a place with different orders. For all of classification the likelihood is figured, after that masterminding is performed with a particular true objective to characterize the tweet as having a place differing computerized risk class. This Framework means to perceive suspicious words from advanced messages and pursue the theorized guilty parties. Starting at now existing Instant Messengers and Social Networking Sites don't have these features of getting significant suspicious instances of danger development from dynamic messages and find associations among people, spots and things in the midst of online visit, as criminals have changed in accordance with it.

RELATED WORK

Mashael Saeed Alqhtani [Base paper], Social networks are the most important communication channels in recent years, which popular among the different social groups. These networks affected the ideas and policies of individuals, groups and communities. Every day, millions of tweets on Twitter are being published. These tweets reflect opinions and beliefs of their publishers and affect others as well. Therefore, it is important to analyze these tweets and identify and classify trends of different users.

This research aims to classify social network to anomaly groups such as: Terrorist and dissident; by analyzing tweets data on the Twitter; then identify an anonymous user’s affiliation to these groups. To address this problem, we first extract a set of features to characterize each group using different data mining techniques and store these features in the database. Text mining, sentiment analysis, and opinion mining techniques will be used to accomplish this extraction. The objective of data extraction is to measure the similarity of selected user tweets with respect to extracted features. It will enable to determine high percentage of similarity between the user tweets and group characteristics to expose his/her affiliation to this group.

SalimAlami, Omar El Beqqali [01], the exponential advancement in information and communication technology has fostered the emergence of new channels for online discussion and has also reduced distances between people. Unfortunately, malicious people take advantage of this technological achievement in the sense that they use it for illegal purposes. In social media, the users produce several and various formats of suspicious posts (text, image, video…) and exchange them online with other people. The data in most social media sites are stored in text format, so in this work we will focus only on text posts. Text mining is an effective way to add semantics aspect to this communication’s form presenting a significant research challenge. Similarity approach is used in text analysis to detect suspicious posts in social media. The evaluation of our proposed approach is done within real posts.

NasiraPerveen Malik M. SaadMissenQaisarRasool [02], Spams are becoming a serious threat for the users of online social networks especially for the ones like of twitter. twitter’s structural features make it more volatile to spam attacks. In this paper, we propose a spam detection approach for twitter based on sentimental features. We perform our experiments on a data collection of 29K tweets with 1K tweets for 29 trending topics of 2012 on twitter. We evaluate the usefulness of our approach by using five classifiers i.e. Bayes Net, Naive Bayes, Random Forest, Support Vector Machine (SVM) and J48. Naive Bayes, Random Forest, J48 and SVM spam detections performance improved with our all proposed features combination. The results demonstrate that proposed features provide better classification accuracy when combined with content and user-oriented features.

AnithaChennamaneni, Shwadhin Sharma, Babita Gupta [03], This exploratory study examines the cyber security attitudes and actual behaviour over time using the data collected on the social media micro blogging platform, Twitter. We plan to use the sentiment analysis and text mining techniques on original tweets related to cyber security collected at two different time periods. Upon completion of this research, we would present the analysis of the relationship between the cyber security attitudes and behaviour and how behaviours may be shaped by the attitudes. This research work aims to contribute to the extant literature in cyber security and endeavours to enhance our understanding of cyber security attitude and behaviour by validating the proposed research model and hypotheses by using real-time, user-generated, social media data.

Carter Chiu, Justin Zhan, and Felix Zhan [04], Multimodal data can be used to gain additional perspective on a phenomenon. For applications, such as security and the detection of suspicious activity, the need to aggregate and analyse data from multiple modes is vital. Recent research in suspicious behaviour detection has introduced methods for identifying and scoring dense blocks in multivariate tensors, which are consistent indicators of suspicious activity. None yet, however, have proposed a method for the merging and analysis of multiple modes of data for suspicious behaviour, especially when the set of items described in each data set do not match that is, the data is partially paired which is common when data sets originate from different sources. Neither has a method been described for dealing with the similar case of incomplete data. This paper introduces a technique for multimodal data analysis for suspicious activity detection when the data are only partially paired and/or incomplete. The method is applied to synthetic and real data, demonstrating strong precision and recall even in poorly paired cases.

Swati Agarwal, AshishSureka [05], Research shows that various social media platforms on Internet such as Twitter, Tumblr (micro-blogging websites), Facebook (a popular social networking website), YouTube (largest video sharing and hosting website), Blogs and discussion forums are being misused by extremist groups for spreading their beliefs and ideologies, promoting radicalization, recruiting members and creating online virtual communities sharing a

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

common agenda. Popular microblogging websites such as Twitter are being used as a real-time platform for information sharing and communication during planning and mobilization if civil unrest related events. Applying social media intelligence for predicting and identifying online radicalization and civil unrest oriented threats is an area that has attracted several researchers' attention over past 10 years. There are several algorithms, techniques and tools that have been proposed in existing literature to counter and combat cyber-extremism and predicting protest related events in much advance. In this paper, we conduct a literature review of all these existing techniques and do a comprehensive analysis to understand state-of-the-art, trends and research gaps. We present a one class classification approach to collect scholarly articles targeting the topics and subtopics of our research scope. We perform characterization, classification and an in-depth meta-analysis. Meta-analysis of about 100 conference and journal papers to gain a better understanding of existing literature.

HarunaIsah, Paul Trundle, Daniel Neagu [06], the growing incidents of counterfeiting and associated economic and health consequences necessitate the development of active surveillance systems capable of producing timely and reliable information for all stake holders in the anti-counterfeiting fight. User generated content from social media platforms can provide early clues about product allergies, adverse events and product counterfeiting. This paper reports a work in progress with contributions including: the development of a framework for gathering and analysing the views and experiences of users of drug and cosmetic products using machine learning, text mining and sentiment analysis; the application of the proposed framework on Facebook comments and data from Twitter for brand analysis, and the description of how to develop a product safety lexicon and training data for modelling a machine learning classifier for drug and cosmetic product sentiment prediction. The initial brand and product comparison results signify the usefulness of text mining and sentiment analysis on social media data while the use of machine learning classifier for predicting the sentiment orientation provides a useful tool for users, product manufacturers, regulatory and enforcement agencies to monitor brand or product sentiment trends in order to act in the event of sudden or significant rise in negative sentiment.

Ali Hasan, Sana Moin, Ahmad Karim and ShahaboddinShamshirband [07], Growth in the area of opinion mining and sentiment analysis has been rapid and aims to explore the opinions or text present on different platforms of social media through machine-learning techniques with sentiment, subjectivity analysis or polarity calculations. Despite the use of various machine-learning techniques and tools for sentiment analysis during elections, there is a dire need for a state-of-the-art approach. To deal with these challenges, the contribution of this paper includes the adoption of a hybrid approach that involves a sentiment analyzer that includes machine learning. Moreover, this paper also provides a comparison of techniques of sentiment analysis in the analysis of political views by applying supervised machine-learning algorithms such as Naïve Bayes and support vector machines (SVM).

Mariam Adedoyin-Olowe, Mohamed MedhatGaber, Frederic Stahl [08], Social network has gained remarkable attention in the last decade. Accessing social network sites such as Twitter, Facebook LinkedIn and Google+ through the internet and the web 2.0 technologies has become more affordable. People are becoming more interested in and relying on social network for information, news and opinion of other users on diverse subject matters. The heavy reliance on social network sites causes them to generate massive data characterised by three computational issues namely; size, noise and dynamism. These issues often make social network data very complex to analyse manually, resulting in the pertinent use of computational means of analysing them. Data mining provides a wide range of techniques for detecting useful knowledge from massive datasets like trends, patterns and rules [44]. Data mining techniques are used for information retrieval, statistical modelling and machine learning. These techniques employ data pre-processing, data analysis, and data interpretation processes in the course of data analysis. This survey discusses different data mining techniques used in mining diverse aspects of the social network over decades going from the historical techniques to the up-to-date models, including our novel technique named TRCM. All the techniques covered in this survey are listed in the Table.1 including the tools employed as well as names of their authors.

PROPOSED METHODOLOGY

In this paper we proposed a framework to process a no of tweets from five different classes and according to the presence of threat work the tweets are classified as suspicious tweets or not and a class level is assigned. The proposed system consist different modules: user interface, text extraction, text pre-processing/cleaning, k-Means clustering and KNN for classification. The real-time tweets from five categories (Murder, Kidnap, Robbery & theft, Corruption Charges and Match fixing) are extracted from twitter and stored as plain text. The proposed system consists of four modules: Tweet Extraction, log pre-processing, Clustering using K-means, and KNN Classification for more accurate categorization of Tweets. This system can classify suspicious or threat tweet more accurate by associating K means with KNN Classification algorithm.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

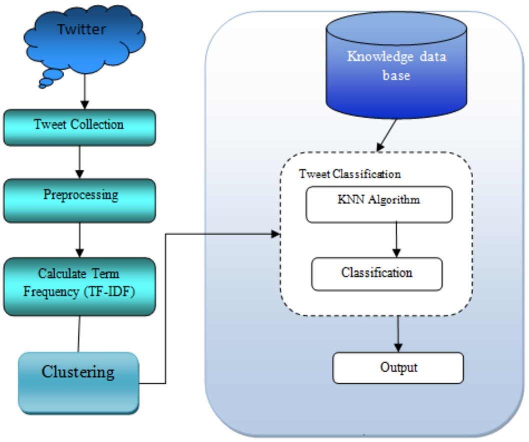


Figure 5.1: - Proposed System Architecture

Extracting tweets from twitter using twitter4j-core-4.0.4 API

The tweets are extricated from every one of five categories. The information recovery is finished by utilizing twitter API twitter4j used to validate the open source system with the twitter application.

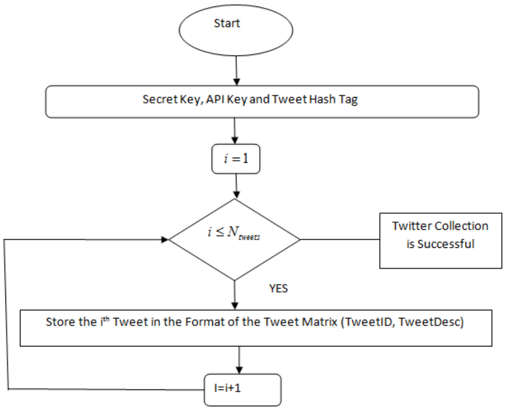

Start
Secret Key, API Key and Tweet Hash Tag
$i = 1$
$i \le N_{tweets}$
Twitter Collection is Successful
YES
Store the ith Tweet in the Format of the Tweet Matrix (TweetID, TweetDesc)
I=i+1

Figure: - process of tweet extraction

Storing tweets

This framework extracts and store tweets from twitter as (TwitterId, TwitterDesc, and UserId). Twitter Id is one of a kind Id related with the tweet, TwitterDesc is the genuine tweet and UserId is the Id related with the client.

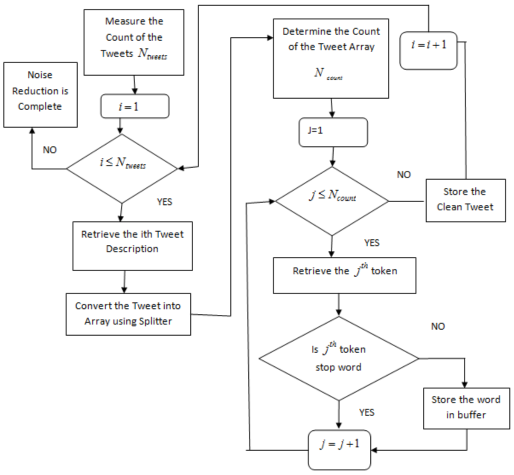
Noise Reduction/ text pre-processing

This process is responsible for removal of noise the stop words that are present in the given tweet.

1. Determine the count of the number of tweets.
2. For each of the tweet the cleaning is performed and the stop words are removed and one can obtain the clean tweet.

Stop words

These are the arrangement of words which don't have a particular importance. The information mining gathering has characterized set of catchphrases. Stop words will be words which are sifted through previously or in the wake of preparing of common dialect information (content). There isn't one distinct rundown of stop words which all frameworks utilize and such a filter isn't generally utilized.


Measure the Count of the Tweets N_{tweets}
Noise Reduction is Complete
$i = 1$
NO
$i \le N_{tweets}$
YES
Retrieve the ith Tweet Description
Convert the Tweet into Array using Splitter
Determine the Count of the Tweet Array N_{count}
$i = i + 1$
J=1
NO
$j \le N_{count}$
Store the Clean Tweet
YES
Retrieve the j^{th} token
NO
Is j^{th} token stop word
Store the word in buffer
YES
$j = j + 1$

K-Means Clustering

K-Means clustering is applied and each tweet is assigned to a cluster based on its closest centroid and the centroid is computed based on mean distance of all tweets weights.

$$\mu(x) = \frac{1}{n}\sum_{k=1}^{n} x$$

Where ***n*** is the no of tweets and the ***x*** corresponds to the each tweet.

KNN Classification

In this process KNN is applied for each tweet from the list of clean tweets.

1. Determine the count of the number of clean tweets.
2. For each of the clean tweet the KNN algorithm is applied and calculates the value for distance metrics of each clean tweet.
3. The value for distance metrics is computed according to the no of threat words present in the each clean tweet.

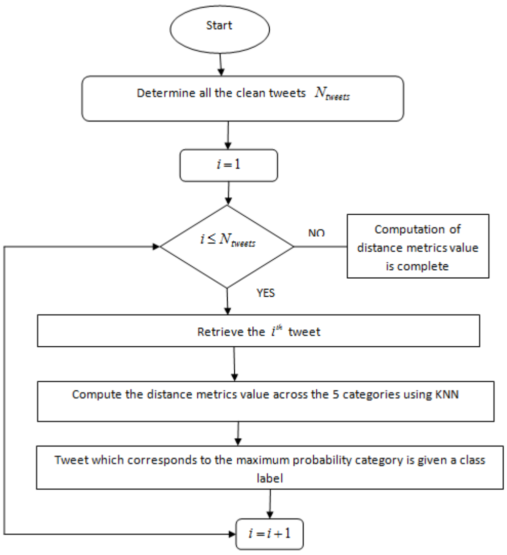

Start
Determine all the clean tweets N_{tweets}
$i = 1$
$i \le N_{tweets}$
NO
Computation of distance metrics value is complete
YES
Retrieve the i^{th} tweet
Compute the distance metrics value across the 5 categories using KNN
Tweet which corresponds to the maximum probability category is given a class label
$i = i + 1$

Figure: - KNN for each tweets

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

After KNN computes value for distance metrics of each tweet for each category, the tweets having highest value of distance metrics is set to as class label.

RESULT ANALYSIS

This section provides results and implementation details of proposed framework.In Figure 4.1 we have presented the clusteringresults obtained after execution of pre-processing. All the clusters are assigned to a nearest centroid.

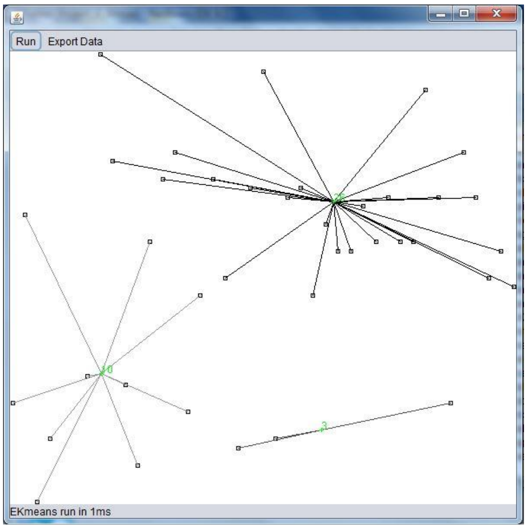


Figure 4.1: - Clustering is performed and each tweet is assigned to closest centroid

Figure 4.2 demonstrates the comparison between accuracy of existing system and proposed system. This result shows that our proposed system is better that existing system by obtaining higher accuracy that existing one.

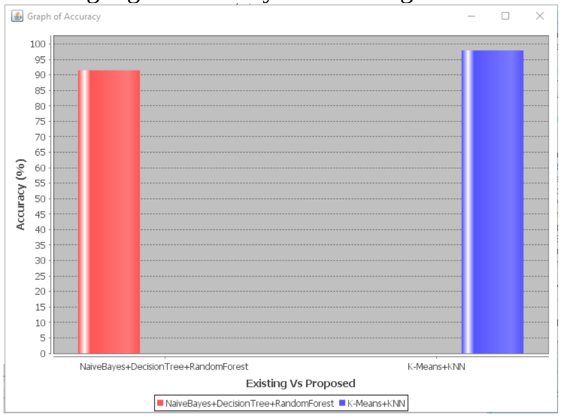


Figure 4.2: - Comparison graph of accuracy between Existing and proposed system

CONCLUSION

This paper proposed an approach to analyse tweets of five threat classes which are Murder, Kidnap, Robbery & theft, Corruption Charges and Match fixing and classify these tweets based on words specified for these classes. Currently existing Instant Messengers and Social Networking Sites lack these features of capturing significant suspicious patterns of threat activity from dynamic messages and find relationships among people, places and things during online chat, as criminals have adapted to it. In this paper we used twitter tweets to detect cyber massages and classify them into five categories based on suspicious words in that tweets. We implement this project on java framework and utilized k-Means and k-NN for classification of tweets. The classification result shows that the proposed system can successfully classify twitter tweets of different classes with better accuracy.

REFERENCES

[1] [Base paper] MashaelSaeedAlqhtani, **"DATA MINING APPROACH FOR CLASSIFYING TWITTER'S USERS",** International Journal of Computer Engineering & Technology (IJCET), Volume 8, Issue 5, Sep-Oct 2017, pp. 42–53, ISSN Print: 0976-6367,

[2] SalimAlami, Omar El Beqqali, **"Detecting Suspicious Profiles Using Text Analysis within Social Media"**, Journal of Theoretical and Applied Information Technology, 31st March 2015. Vol.73 No.3, ISSN: 1992-8645,

[3] NasiraPerveen Malik M. SaadMissenQaisarRasool, **"Sentiment Based Twitter Spam Detection",** (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 7, 2016, pp:568-573,

[4] AnithaChennamaneni, Shwadhin Sharma, Babita Gupta, **"Twitter Sentiment Analysis: An Examination of Cybersecurity Attitudes and Behavior"**, Proceedings of the 2016 Pre-ICIS SIGDSA/IFIP, Special Interest Group on Decision Support and Analytics (SIGDSA), 2016, http://aisel.aisnet.org/sigdsa2016/17,

[5] Carter Chiu, Justin Zhan, and Felix Zhan, **"Uncovering Suspicious Activity from Partially Paired and Incomplete Multimodal Data"**, IEEE Access (Volume: 5), Page(s): 13689 – 13698, 12 July 2017, DOI: 10.1109/ACCESS.2017.2726078, ISSN: 2169-3536,

[6] Swati Agarwal, AshishSureka, **"Applying Social Media Intelligence for Predicting and Identifying On-line Radicalization and Civil Unrest Oriented Threats"**, https://arxiv.org/abs/1511.06858v1, 21 Nov 2015, pages: 01-18,

[7] HarunaIsah, Paul Trundle, Daniel Neagu, **"Social Media Analysis for Product Safety using Text Mining and Sentiment Analysis"**, 2014 14th UK Workshop on Computational Intelligence (UKCI), 20 October 2014, DOI: 10.1109/UKCI.2014.6930158, ISSN: 2162-7657,

[8] Ali Hasan, Sana Moin, Ahmad Karim and ShahaboddinShamshirband, **"Machine Learning-Based Sentiment Analysis for Twitter Accounts",** Applied Modern Mathematics in Complex Networks, 24 February 2018, https://doi.org/10.3390/mca23010011,

[9] Mariam Adedoyin-Olowe, Mohamed MedhatGaber, Frederic Stahl, **"A Survey of Data Mining Techniques for Social Network Analysis",** Journal of Data Mining & Digital Humanities, 2014 (June 24, 2014), https://arxiv.org/abs/1312.4617,

[10] Sudip Mittal, Prajit Kumar Das, VarishMulwad, **"Cyber-Twitter: Using Twitter to generate alerts for Cyber-security Threats and Vulnerabilities"**, 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM),

[11] SanghoLeey and Jong Kim, **"WARNINGBIRD: Detecting Suspicious URLs in Twitter Stream"**, IEEE Transactions on Dependable and Secure Computing (Volume: 10, Issue: 3, May-June 2013), DOI: 10.1109/TDSC.2013.3, Page(s): 183 – 195, 11 January 2013, ISSN: 1545-5971,

[12] Mercy Paul Selvan and RenukaSelvaraj, **"Mining user Message Pattern for Suspicious Behaviour on Terrorism using NLP in Social Networks with Single Sign-On",** Indian Journal of Science and Technology, Volume 10(14), April 2017, ISSN: 0974-6846, DOI: 10.17485/ijst/2017/v10i14/111364,

[13] Ameena A, **"Detection of User Cluster with Suspicious Activity in Social Networking Sites using Natural Language Processing"**, International Journal of Innovative Research in Computer and Communication Engineering, Volume 3, Issue 10,

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

October 2015, DOI: 10.15680/IJIRCCE.2015. 0310052, pages: 9483-9489

[14] Matthew S. Gerber, **“Predicting Crime Using Twitter and Kernel Density Estimation”,** ELSVIER, Decision Support Systems, Volume 61, May 2014, Pages 115-125, https://doi.org/10.1016/j.dss.2014.02.003,

[15] Anna Sapienza, Alessandro Bessi, SaranyaDamodaran, Paulo Shakarian, Kristina Lerman, Emilio Ferrara, **“Early Warnings of Cyber Threats in Online Discussions”**, 2017 IEEE International Conference on Data Mining Workshops (ICDMW), pp:667-674, 2017, ISSN: 2375-9259, https://doi.org/10.1109/ICDMW.2017.94,

[16] Daniel E. O’leary, **“Twitter Mining for Discovery, Prediction and Causality: Applications and Methodologies”**, International Journal of Intelligent Systems in Accounting and Finance Management, https://doi.org/10.1002/isaf.1376, Volume 22 Issue 3, July 2015, Pages 227-247,

(This work is licensed under a Creative Commons Attribution 4.0 International License.)