Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 5, Issue 02, February -2018

A SURVEY PAPER ON PREDICTION OF ONLINE SPREAD OF TERRORISM ON TWITTER

Mr Anilkumar Munani[1], Mr Bhavesh Tanawala[2], Mr Prashant Swadas[3]

1Computer Department, BVM Engineering College, V.V.Nagar, Gujarat, India
2 Computer Department, BVM Engineering College, V.V.Nagar, Gujarat, India
3 Computer Department, BVM Engineering College, V.V.Nagar, Gujarat, India

Abstract — *Terrorist groups like al-quida, Indian mujahedeen, ISIS and other terrorist groups are spreading their propaganda using internet or different social media websites like facebook, Twitter and Google+. Basic idea to stop or reduce spreading of terrorism is to remove all this accounts. To implement this idea needs lots of human efforts which includes reading lot of information and analyzing contain. So to reduce human efforts we will make a system which detect message given by terrorist group on twitter. Our system will classify tweets and finds tweets are supporting ISIS group or not. We want to build a system which will give better result for analyzers.*

Keywords- *Terrorism, ISIS, social media radical content, text mining, natural language processing, user cluster*

I. INTRODUCTION

Today, terrorist groups are well formed with so many resources; social media such as twitter is one of them. Many radical groups use their online twitter accounts to spread their propaganda. Terrorist groups use their network organizations and strategies to put audio video or text content [4]. So, it has become very important to control terrorism and stop their spread before certain amount of time [1]. As observed some previous work, they have proposed some efficient algorithm or designed a system to capture radical cluster or radical content on twitter.

II. RELATED WORKS AND LITERATURE SURVEY

In [4], authors have used machine learning approach to classify tweets. They have build classifier on the basis of three dataset.

Dataset	Description
TW-PRO	Tweets that are pro-ISIS
TW-RAND	Randomly collected tweets
TW-CON	Tweets that are against ISIS

Table I : dataset used for experiment in [4] .

Author have focused on the English hash tag which are #IS, #ISLAMICSTATE, #LOVEISIS, #ALLEYESONISIS and other hash tags [4]. Authors have used three different classes of features.

A. STYLOMETRIC FEATURE

Stylometric feature contains most frequent used words like state, Islamic, not, do, kill , support, abu, allah, people and al.

Function words	Frequency of various function words	293
Frequent words	Frequency of major frequently used words	173
Punctuation	Frequency of characters,, . , [,] , ! , ? , &	13
Hash tags	Frequency of most frequent letter bigrams	100
Letter bigrams	Frequency of most frequent letter bigrams	133
Word bigrams	Frequency of most frequent words bigrams	99

Table II The list of words that have been used in [4].

Stylometric features also include punctuation, letter bigrams, word bigrams and the most frequently used hash tags [4].

B. TIME BASED FEATURE

Time based feature contain detailed description about when tweet is posted . The following attributes are specified in [4]:

• Hour Of Day: Hour1, Hour2, . . . , Hour24,

• Period Of Day: Morning, Afternoon, Evening, Night, Mid Night.

• Day: Sunday, Monday, . . . , Saturday
•Type Of Day: WeekDay, WeekEnd.

C. SENTIMENT BASED FEATURE

Sentiment analysis determines the attitude of text towards a specific topic. The analysis of sentiment was done using natural language processing, the values the sentiment can take are: very negative, negative, neutral, positive, very positive. Authors have used three different classifiers AdaBoost, Naïve Bayes and SVM using all features on all datasets [4]. They have finally got result that AdaBoost performs slightly better than both Naïve Bayes and SVM.

To detect terrorism, authors in [1] have designed terrorism analysis system using techniques like DOM (Document Object Model) tree for web data extraction, SVM algorithm for classification; SIFT algorithm for object recognition and extraction and k-mean algorithm for segmentation of textual data.

Overall goal of [2] is to develop a dynamic online mechanism that capture and counter violent extremist narratives or terrorist who are using social media. To achieve this they mostly use big data analytics and also use a variety of computational or mathematical methods like Natural Language Processor, Semantic Web and Crowd Sourcing [2].

The social network context provides a set of methods to analyze the structure of whole social entity as well as a variety of theoretical information explaining the patterns observed in these structures [3]. Similarly terrorist large networks have an identical structure like social networks where each node is directly or indirectly linked to the terrorist organization [3].

Advanced Terrorist Detection System (ATDS) is aimed at tracking down online access to radical content [5]. ADTS operates in two modes :
1) training mode 2) detection mode [5]
In training mode author have designed terrorist transaction database acknowledge their behavior from their internet activity. In detection mode author have calculated threshold value and content based detection [5]. Author had extracted data from log files over the net and check whether those URL's are suspicious or not. Following are ways:
1) Check Browser's History
2) Crawl's URL's
3) Text mining from web pages

The actions of malicious users can harm both social media users and those who want to use information in social media to understand of culture pulse of a specific area [6].
Using large dataset containing data from multiple countries over multiple months, we find that the removal of suspended users can have profound impacts on what users are defined as influential, the overall topology of methods and co-topic network and less impact on the identification of what being tweeted about [6]. Author analyzed network analysis and spam detection [6].

Author in [7], highlighted that 90% of terrorist activities carried out on the internet are organized through social networking sites. They have designated a system which is not fully automated to achieve this system requires derivative user involvement [7]. Their proposed system acquires input through a fixed database of e-mail [7]. This e-mail is drawn from the Enron e-mail dataset. They proposed a system design which identifies the cluster of people or radical groups in social networking sites, whose behavior are suspicious [7]. They have also focused on finding the cluster of users who are discussing about same topic, which is done by finding similarities in message which are being exchanged among social media users. The design of author's proposed system is collection of five sub system [7]. Which are following:
• Online data monitoring system and Database
• Suspicious message identification using
NLP/Keyword system
• Latent semantic analysis (LSA) system
• Suspicious users identification system
• Visual representation of suspicious users.

Author in [8] applied their Algorithm on the expanded data from facebook Operation [8]. The algorithm detected active systems or nodes that can recruits other nodes in the group easily because of them position among other nodes in the network [8]. Detecting the active nodes is done by using combination of different popular centrality measures on the nodes in the group [8].

The Dark Web Forum Portal (DWFP) maintains a collection of 29 online jihadist forums, which currently contains 14,297,961 messages and 1,553,122 threads from 362,495 authors [9]. They had discussed only two dark web forums Islamic Network and Islamic Awaking Forum and their data has been retrieved from dark web portal [9].
By using dark web analysis, security agencies can execute vigilance and data collection for CT because dark webs are large source of information. This analysis will help security agencies to detect and avoid terrorist threats. Dark web analysis will uncover the hidden patterns and connect the dots in information space. Thus, Dark web analysis can be used for detecting and avoiding terror threats or radical activity.

III RESEARCH GAP

In [4] many of the features are dependent on the dataset. We can use both data dependent and data independent features and evaluate the result. In [1] author have used DOM tree to extract information from the web to do web mining. In [2], [3], [7] and [8] they have focused on to find cluster of people where suspicious activity found. In [5] author have created own system, but not used specific algorithm for accurate results. In [9] they have used dark forum portals and analyzed terrorist activities.

REFERENCES

[1] Ms. Pooja S. Kade1, Prof. N.M. Dhande, " A Paper on Web Data Segmentation for Terrorism Detection using Named Entity Recognition Technique" presented at International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056,Volume: 04 Issue: 01 | Jan -2017,

[2] Budak Arpinar & Ugur Kursuncu and Dilshod Achilov, "Social Media Analytics to Identify and Counter Islamist Extremism: Systematic Detection, Evaluation, and Challenging of Extremist Narratives Online" presented at International Conference on Collaboration Technologies and Systems. 978-1-5090-2300-4/16 2016 IEEE.

[3] Surajit Dasgupta, Chandan Prakash, "Intelligent Detection of Influential Nodes in Networks" presented at International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) , 978-1-4673-9939-5/16, 2016-IEEE,

[4] Michael Ashcroft, Ali Fisher, Lisa Kaati, Enghin Omer, Nico Prucha, "Detecting Jihadist Messages on Twitter" presented at European Intelligence and Security Informatics Conference, 978-1-4799-8657-6/15, 2015 IEEE.

[5] Sonali Vighne, Priyanka Trimbake, Anjali Musmade, Ashwini Merukar, Sandip Pandit, "An Approach to Detect Terror Related Activities on Net" presented at IJARIIE-ISSN(O)-2395-4396, Vol-2 Issue-1 2016.

[6] Wei Wei Carnegie, Kenneth Joseph, Huan Liu, Kathleen M. Carley," The Fragility of Twitter Social Networks Against Suspended Users" International Conference on Advances in Social Networks Analysis and Mining, 2015 IEEE/ACM.

[7] Sharath Kumar A and Sanjay Singh, " Detection of User Cluster with Suspicious Activity in Online Social Networking Sites" Second International Conference on Advanced Computing, Networking and Security, 978-0-7695-5127-2/13,2013 IEEE,.

[8] Ala Berzinji, Frzand Sherko Abdullah, Ali Hayder kakei, "Analysis of Terrorist Groups on Facebook", 978-0-7695-5062-6/13,IEEE.

[9] Abhishek sachan, "Countering Terrorism through Dark Web Analysis" ICCCNT'12, 26th _28th July 2012, Coimbatore, India, IEEE-20180.

