Open Source Social Media Analytics for Intelligence and Security Informatics Applications

Swati Agarwal[1], Ashish Sureka[2(✉)], and Vikram Goyal[1]

[1] Indraprastha Institute of Information Technology, Delhi (IIITD), New Delhi, India
{swatia,vikram}@iiitd.ac.in
http://www.swati-agarwal.in, http://www.iiitd.edu.in/~vikram/
[2] Software Analytics Research Lab (SARL), New Delhi, India
ashish@iiitd.ac.in
http://www.software-analytics.in/

Abstract. Open-Source Intelligence (OSINT) is intelligence collected and inferred from publicly available and overt sources of information. Open-Source social media intelligence is a sub-field within OSINT with a focus on extracting insights from publicly available data in Web 2.0 platforms like Twitter (micro-blogging website), YouTube (video-sharing website) and Facebook (social-networking website). In this paper, we present an overview of Intelligence and Security Informatics (ISI) applications in the domain of open-source social media intelligence. We present technical challenges and introduce basic Machine Learning based framework, tools and techniques within the context of open-source social media intelligence using two case-studies. The focus of the paper is on mining free-form textual content present in social media websites. In particular we describe two important application: online radicalization and civil unrest. In addition to covering basic concepts and applications, we discuss open research problem, important papers, publication venues, research results and future directions.

Keywords: Information retrieval · Intelligence and Security Informatics · Machine learning · Mining user generated content · Open-Source Intelligence · Social media analytics

1 Introduction

Terrorism is not only a national problem specific to few countries experiencing terrorist incidents but an international and a global problem. Countering terrorism and building advanced technology based solutions to combat terrorism is a major concern and challenge to the government, law enforcement agencies and society [11,16]. Recent research and evidences demonstrate that Internet and social media platforms are increasing used by extremists and terrorists for planning and mobilization, recruitment, propagating hate and disseminating extreme political and religious beliefs. Social media platforms (such as YouTube

N. Kumar and V. Bhatnagar (Eds.): BDA 2015, LNCS 9498, pp. 21–37, 2015.
DOI: 10.1007/978-3-319-27057-9_2

video-sharing website and Twitter micro-blogging platform) are exploited for conducting extremist and terrorist activities due to low content publication barrier, wide reachability to a large number of people across countries and anonymity. Online radicalization has a major impact on society that contributes to the crime against humanity and main stream morality. Presence of such content in large amount on social media is a concern for website moderators (to uphold the reputation of the website), government and law enforcement agencies (locating such users and communities to stop hate promotion and maintaining peace in country). Hence, automatic detection and analysis of radicalizing content and protest planning on social media are two of the important research problems in the domain of ISI [4,10]. Monitoring the presence of such content on social media and keeping a track of this information in real time is important for security analysts working for law enforcement agencies. Mining publicly available open-source social media data for analysing, detecting, forecasting and conducting a root-cause analysis is an area that has attracted several researchers' attention [2,15] as shown in Fig. 1, the focus and scope of this paper is an intersection of three fields: (1) Online Social Media Platforms, (2) Intelligence and Security Informatics, and (3) Text Mining and Analytics. Online radicalization detection and civil unrest forecasting from social media data are subtopics (focus of this paper) within the broad area of Intelligence and Security Informatics. In this paper, we introduce the topic of online radicalization detection and civil unrest prediction, present technical challenges, problems and solution framework, literature survey, reference to case-studies and future research challenges.

Basic Definitions and Background Concepts

1.1 Online Social Media

Online Social Media Online Social Media and Web 2.0 (also referred to as the current generation of Web) consists of websites such as YouTube (video-sharing), Twitter and Tumblr (micro-blogging), Facebook (social networking), StackOverflow (community based question and answering), Delicious (social bookmarking), online wikis, message boards and discussion forums. Social media platforms are highly participatory and collaborative in nature in which users can easily share content and post messages and comments. According to information on Twitter website[1], there are about 500 million Tweets posted every day and the micro-blogging platform has 316 million monthly active users. According to information on YouTube website[2], YouTube has more than 1 billion users and 300 hours of video to the video-sharing platform every minute. According to SocialMedia-Today August 2015 statistics, 1.925 Billion users utilise their mobiles for Social Media platforms[3].

[1] https://about.twitter.com/company.

[2] https://www.youtube.com/yt/press/statistics.html.

[3] http://www.socialmediatoday.com/social-networks/kadie-regan/2015-08-10/10-amazing-social-media-growth-stats-2015.

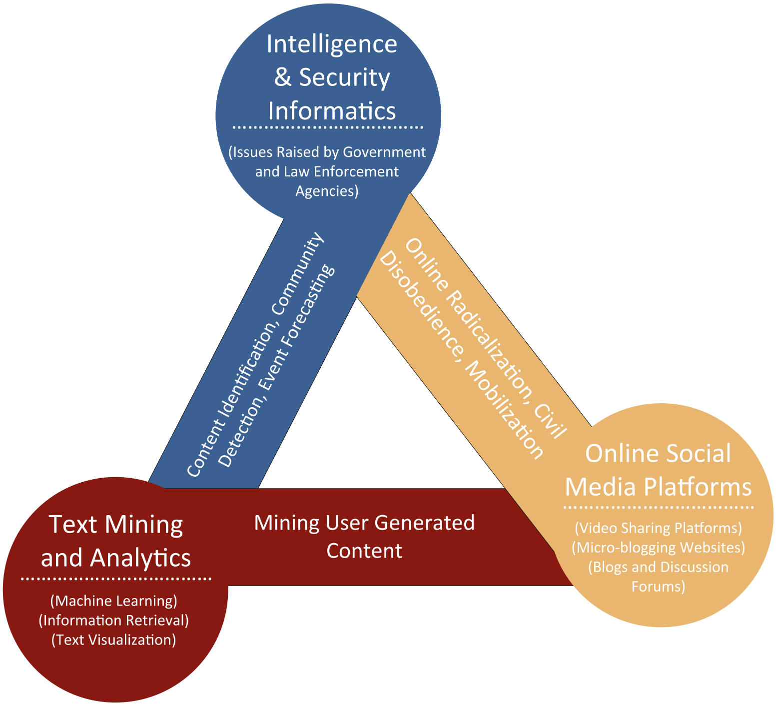

Fig. 1. Scope and focus of research area: an intersection of 3 topics and sub-topics (Online social media, text mining & analytics and Intelligence & Security Informatics)

1.2 Intelligence and Security Informatics

Intelligence and Security Informatics (ISI) is a field of study concerning investigation and development of counter terrorism, national and international security support systems and applications. ISI is a fast growing interdisciplinary area which has attracted several researchers and practitioners (form academia, government and industry) attention spanning across multiple disciplines like computer science, social sciences, political science, law and even medical sciences [16]. An example of ISI application is mining open-source (publicly available) social media data for terrorism forecasting or terrorism root cause analysis [13].

1.2.1 Online Radicalization

Online radicalization consists of using Internet as a medium by extremists and terrorists for conducting malicious activities such as promoting extreme social, religious and political beliefs, recruitment of youth, propagating hatred, forming communities, planning and mobilizing attacks [5–7, 13]. Creating channels and video clips on how to make a bomb or a speech on racist propaganda and posting it on a popular video sharing website like YouTube falls under online radicalization [2, 4, 11]. Using a widely used micro-blogging platform such as Twitter

Fig. 2. Top 3 Twitter accounts posting extremist content on website

for planning and mobilization of terrorist attacks and criminal activities is an example of using Internet and online social media for radicalization [3]. Similarly, posting hate-promoting blogs and comments on online discussion forums to disseminate extreme religious beliefs is a major online radicalization concern for government and law enforcement agencies [16]. Figure 2 illustrates top 3 Twitter accounts of hate promoting users posting extremist content very frequently and having a very large number of followers[4]. According to SwarmCast journal article and Shumukh-al-islam posts, these are the three most important jihadi and support sites for Jihad and Mujahideen Twitter[5].

1.2.2 Civil Unrest and Disorder

Civil unrest is referred to as social instability and protest movements at the National and International level primarily against the government and policy-makers [12]. Civil unrest can be both non-violent demonstrations or strikes as well violent riots. The reason behind large scale civil unrest is mainly discontent in the society due to poor social and economic conditions [15]. The Arab Spring democratic uprising which originated in Tunisia in December 2010[6] and then propagated across various countries in the Arab world in the year 2011 is an example of intense civil unrest and disorder[7].

2 Technical and Computational Challenges

1. **Massive Size and Rich User Interaction.** The volume and variety of data in-terms of the modality such as free-form text, images, audio and video generated every day is so huge (refer to some statistics given in Sect. 1.1) that it poses hard computational challenges such as data processing and storage for researchers or application developers interested in analyzing the data [3,12]. In addition to the massive data size, the rich user interaction (such as friend, follower, subscriber, favorite and like relationship between various users) possible in social media increases data complexity, dimensions and variability which needs to be addressed.

---

4 http://www.terrorismanalysts.com/pt/index.php/pot/article/view/426/html.

5 http://jihadintel.meforum.org/identifier/149/shumukh-al-islam-forum.

6 http://middleeast.about.com/od/humanrightsdemocracy/tp/Arab-Spring-Uprisings.htm.

7 http://www.npr.org/2011/12/17/143897126/the-arab-spring-a-year-of-revolution.

Fig. 3. Examples of multi-lingual posts on extremism and civil unrest

2. **High Velocity.** In addition to the overall quantity, the speed (for example few megabytes per second or minute or millions of rows per hour) at which data is generated poses data analytics challenges [1]. Real-time processing and storing of such high-velocity and massive flow of data is computationally challenging from the perspective of data ingestion [2].
3. **Multilingualism.** The web and social media is inherently multilingual. Content is posted in several different languages and processing them through automated algorithms requires linguistic resources for each language. Mining multilingual information is important and even essential for several Intelligence and Security Informatics (ISI) based application due the global diversity, user demographics and reach of the social media platforms and its users. Figure 3 shows examples of extremism and unrest related tweets posted in Arabic and Dutch language respectively.
4. **Noisy Content.** A huge amount of content on social media is of low quality (such as spam) and in general of low relevance (such as posting a message on what one had for breakfast or lunch) due to low barrier to publication [1,5,8]. Moreover, there are several issues such as grammatical mistakes, spelling errors, usage of non-standard acronyms and abbreviations, emoticons and incorrect capitalization of terms due to the informal nature of the content.
5. **Spam and Fake Accounts.** Spam, irrelevant and unsolicited messages as well as fake accounts is common in social media platforms [8]. Such content not only decreases the value of the website and the user experience but also poses technical challenges for social computing researchers in terms of data cleaning and pre-processing before building accurate data mining models.
6. **Data Annotation and Ground Truth.** Data annotation and ground-truth creation is a basis for several machine learning tasks such as predictive modeling and classification. Creating and annotating high quality ground-truth data at a large scale is effort intensive if done manually and a non-trivial technical challenge if done (semi)-automatically [3].
7. **Manipulation, Fabrication and Adversarial Behavior.** Deception, manipulation, misinformation, adversarial behaviors and credibility is a major issue in social media. Research shows that lot of content posted on social

media is factually incorrect and is rumor [14]. Fake information, rumors and manipulated content is not only a social media abuse but a challenges for researchers and developers in building computational frameworks for Intelligence and Security Informatics (ISI) based applications.

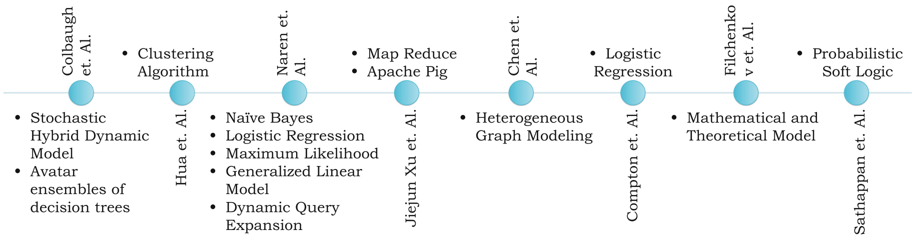

Fig. 4. Machine learning techniques used by researchers in existing literature of event forecasting in the domain of civil disobedience

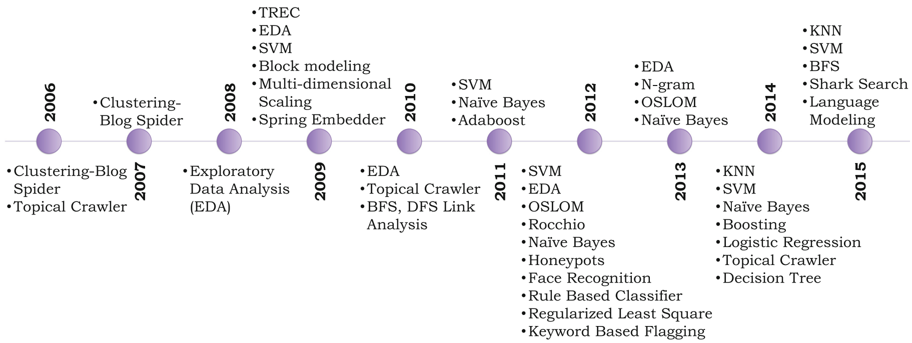

Fig. 5. Machine learning techniques used by researchers in existing literature of online radicalization and hate promotion detection

3 Machine Learning and Data Mining Techniques

In this Section, we discuss the most popular and common data mining techniques used by previous researchers in the area of detecting online radicalization and predicting civil unrest related events. Figure 4 illustrates a preview of these techniques used in existing literature in the domain of civil unrest event detection. Similarly, Fig. 5 shows a timeline of online radicalization techniques proposed over the past decade.

Based on our literature survey and analysis of previous work, we observe that ensemble learning on various machine learning techniques is a common approach to achieve better accuracy in the results. However, Ramakrishnan et al. [15]

proposed a different technique for different type of dataset/activities. They propose 5 different techniques for predicting events from 5 different models: volume based, opinion based, tracking of activities, distribution of events and cause of protest. In ensemble learning approach, Clustering, Logistic Regression and Dynamic Query Expansion (DQE) are the most commonly used techniques. We also observe that Named Entity Recognition (NER) is the most widely used phase in all techniques illustrated in Fig. 4. In named entity recognition, they extract various named entities present in the contextual metadata (tweets, Facebook comment, News article). For example, spatiotemporal expressions (locations and time related enetities), topic being discussed in the tweets or articles. They further use Dynamic Query Expansion and semi-automatic approaches to expand the list of extracted entities and keywords. They use various lexical sources (for example- WordNet[8], VerbOcean[9]) to extract similar and relevant keywords for their case study. Dynamic Query Expansion is an iterative process and converges once same keywords starts repeating. They further perform various clustering, regression (Logistic Regression) and classification (Naive Bayes) algorithm on these entities for event prediction.

We conducted a literature survey in the domain of online radicalization (extremist content, users and communities) detection on social media websites (video sharing websites, micro-blogging websites, blogs and forums) and we collected 37 articles (journal and conference proceedings). Therefore, we create a timeline of various deradicalization approaches that has been used by researchers over the past decade (refer to Fig. 5). Nowadays, social networking websites (Twitter, YouTube, Tumblr etc.) are some of the largest repositories of user generated content (contextual data) on web. Therefore, Text Classification KNN, Naive Bayes, SVM, Rule Based Classifier, Decision Tree, Clustering (Blog Spider), Exploratory Data Analysis (EDA) and Keyword Based Flagging (KBF) are the most commonly used techniques to identify hate promoting content on Internet. However, Topical Crawler and Link Analysis (Breadth First Search, Depth First Search, Best First Search) are common techniques to navigate through links and locate users with similar interest. Language modeling, n-gram and Boosting are other techniques used for mining textual data and classifying hate promoting content on web (social media, blogs, forums) [2]. Figure 5 also reveals that with the emergence of social media websites, Social Networking Analysis (SNA) became a popular techniques among researchers to locate hidden communities and groups. As we looked into the literature, we find that OSLOM (Order Statistics Local Optimization Method) is another common technique that has been used to locate hidden communities of extremist users. It is an open source visualization tool using a clustering algorithm designed to compute the similarity between two nodes (users) and locate their groups (network) sharing a common agenda. OSLOM uses an iterative process to identify internal structure of clusters and overlapping clusters (if possible). In order to detect hidden communities, it locally optimize the clusters by taking edge directions, edge weight, overlapping clusters and hierarchies of clusters into account.

[8] https://wordnet.princeton.edu.

[9] http://demo.patrickpantel.com/demos/verbocean/.

4 Case Studies

In this Section, we present two case studies in the area of Intelligence and Security Informatics. We describe our proposed solution approaches and algorithms for detecting online radicalization on YouTube and predicting civil unrest related events using an open source Twitter data.

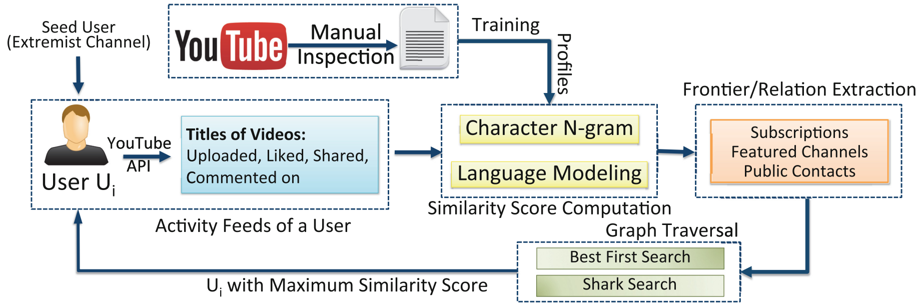

Fig. 6. A focused Crawler based framework for detecting hate promoting videos, users and communities on YouTube

4.1 Identification of Extremist Content, Users and Communities

Research shows that YouTube has become a convenient platform for many extremist groups to post negative videos disseminating hatred against a particular religion, country or person. These groups put forth hateful speech, offensive comments and promote certain ideology and beliefs among their viewers [2]. Social networking allows these extremist users to connect with other hate promoting users and groups, spreading extremist content, facilitating recruitment and forming their virtual communities [4]. Due to the dynamic nature of website, it is technically challenging and overwhelmingly impractical to detect such hate promoting videos and users by keyword based search. In this case study, we formulate our problem of identifying extremist content as a classification problem and use topical crawler based approach to locate such users and their hidden communities sharing a common agenda.

Figure 6 presents a general research framework for the proposed solution approach. The proposed method is a multi-step process primarily consisting of three phases, collection of training profiles, learning the statistical model and topical crawler. We perform a manual inspection on various YouTube channels and collect 35 hate promoting user profiles based upon their activity feeds such as titles and content of videos uploaded, commented, liked and favorited by the user. We build our training dataset by extracting contextual metadata of all activity feeds and profile summary of these 35 channels using YouTube API[10]. We also perform a characterization study on the contextual metadata and observe some

[10] https://developers.google.com/youtube/getting_started.

domain specific key-terms. We divide these terms into following 9 categories: important dates, religion, region, people name, negative emotions, communities, politics terms, war related terms and others. We apply character n-gram and language modeling on our training dataset and build our learning model. We use this learning model to compute the extent of similarity between contextual metadata of input channel and training data. We classify a YouTube channel as relevant (hate promoting) or irrelevant if the similarity score is above or below a threshold value. Further, To mine the relations among these hate promoting users, we build a focused crawler that follows the classifier and extract external links (connected users) to that profile. This focused crawler takes one YouTube channel (annotated as hate promoting channel) as a seed and classifies it using the learning model. We further extract the frontiers (featured channels, public contacts and subscribers) of this channel by parsing channel's YouTube homepage using jsoup HTML parser library[11]. We compute a similarity score for each frontier against training data and prioritize them in the descending order of their score. We execute our focused crawler for each frontier that results into a directed connected graph. In this graph, nodes represents the YouTube channel and edges represents the link between two channels (e.g. features channel). We further perform social network analysis on this graph to locate hidden communities and influential users playing major role in the community. As shown in the Fig. 6, in graph traversal phase, we use two different algorithms: Best First Search (BFS) and Shark Search (SS) traversal. In BFS algorithm, we extract the frontiers of a user channel only if it is classified as a relevant (hate promoting) channel. While in Shark Search algorithm, we explore the frontiers of both relevant and irrelevant users. Given a graph G, if B is a frontier of A and A is classified as irrelevant then in SS Algorithm, the similarity score of the B is reduced by a decay factor d that directly impacts on the priority of user. Therefore a frontier node of an irrelevant node has an inherited score of $score_{frontier} * d$. However, this inherited score is dynamic because as we traverse in the network, a node might have more than one parent (frontier of multiple channels). If a node has multiple parent nodes then we chose the $max(S_{P1}, S_{P2} ..... S_{Pn},)$ where S_{Pi} is the inherited score of frontier for i_{th} parent.

Figure 7 (generated using ORA[12]) shows the cluster representation of network graph using Best First Search and Shark Search focused crawlers. Colors of nodes represents the variation in number of frontiers per user and the width of the edge represents the number of links between two users. Based on the network measurements, we find that the average clustering coefficient of graph for shark search focused crawler is very large in comparison to the graph for BFS focused crawler. Similarly, Fig. 7 also reveals that in BFS approach, network has 13 different clusters (represented as 13 different colors) for a total of 23 nodes. Among these 13 clusters, 6 clusters consists of only one user node which shows the difference in user channels. However, the cluster representation of network for shark search focused crawler contains only 7 clusters for 24 nodes. Unlike

[11] http://jsoup.org/apidocs/.
[12] http://www.casos.cs.cmu.edu/projects/ora/.

BFS focused crawler, in Shark search approach, only 2 clusters are formed with one node and we are able to find 3 existing strongly connected communities: C, D, E, H, I and B, C, D, E, G, H, I, J and A, D, E, G, where G has the maximum closeness centrality and connected to maximum number of users.

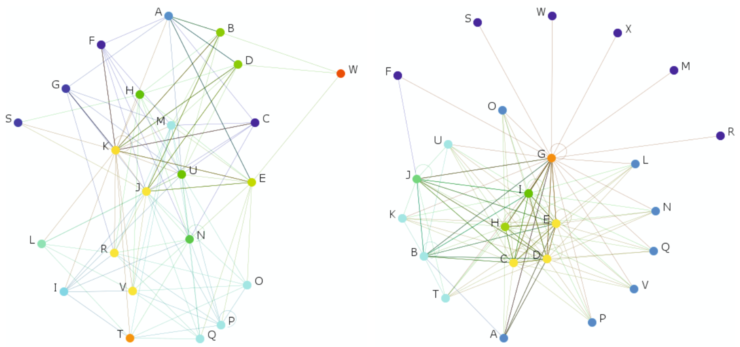

Fig. 7. Cluster graph representation for best first search (Left) Shark search Crawler (Right). Source: Agarwal et al. [4] (Color figure online)

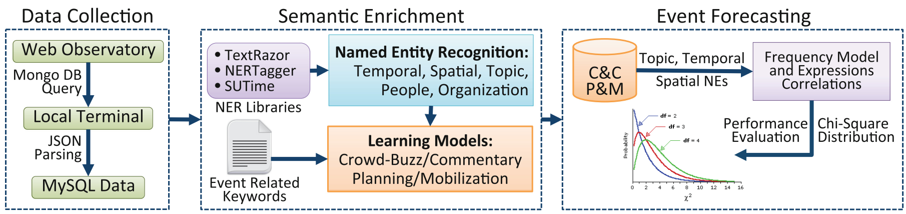

Fig. 8. A general research framework for mining Tweets and predicting civil unrest related events

4.2 Event Forecasting for Civil Unrest Related Events

We define the problem of civil unrest event prediction as the identification of several named entities expressions: spatial- location l where the protest is going to happen, temporal- time ti (can be date and time) of the protest and topic to- root cause of the protest. In this case study, we divide our problem into two sub-problems: (1) to perform a characterization study on open source Twitter dataset to investigate the feasibility of building event forecasting model. and (2) to evaluate the performance of machine learning and statistical based forecasting model by conducting experiments on real word Twitter data. Figure 8 presents the general research framework of our proposed approach for civil unrest event forecasting on micro-blogs. Figure 8 shows a high level architecture and design of proposed methodology that primarily consists of three phases: data collection, semantic enrichment and event forecasting. We conduct experiments on open source Twitter dataset ('Immigration Tweets') downloaded from Southampton Web Observatory[13]. This dataset consists of approximately $2M$ tweets where each tweet contains at least of these 4 words: 'immigration', 'immigrant', 'migration' and 'migrant'. These tweets are collected in a time span of 5 months collected from October 1, 2013 to February 28, 2014. We use mongoDB queries to access this data on our local terminal and further use JSON parsing to convert it into mySQL. Since this dataset contains all tweets related to 'immigration', we search popular media and news articles posted during this time span and look for immigration related events happened worldwide. We find 5 such civil unrest related events. However, in this paper we present a case study for only one event i.e. 'Christmas Island Hunger Strike'. This event happened in Australia on January 16, 2014 where nine asylum seekers stitch up their mouth with dental floss and protested against new immigration law proposed by the country.

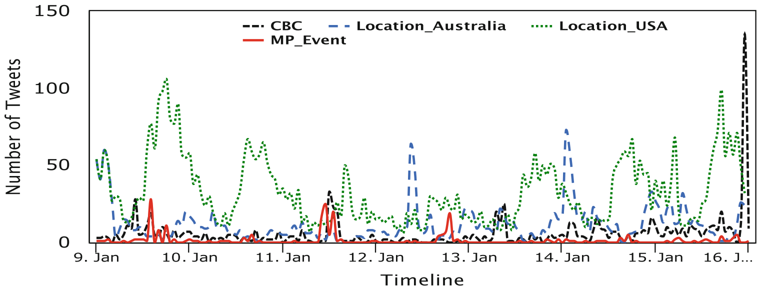

Fig. 9. Trend of crowd-buzz, mobilization and location based Tweets

As shown in the Fig. 8, we perform pre-processing on dataset and semantically enrich these tweets to achieve better accuracy in prediction. For semantic enrichment of tweets, we use a trend based sliding window frame of 7 days and extract only those tweets from the dataset that were published during that time frame. For example, for 'Christmas Island Hunger Strike', we use tweets posted from January 9 to 15. We make a hypothesis that before an event happens there are two types of tweets that are posted: people spreading the word and discussing about the event (crowd-buzz and commentary) and organization posting tweets for planning and mobilizing that event. Based on this hypothesis, we train a multi-class classifier that classifies tweets into 3 categories: Crowd-buzz and Commentary (C&C), Planning and Mobilization (P&M), Irrelevant (not about the event). We use C&C and P&M tweets as two lead indicators that helps in early detection of an event. These indicators filters relevant tweets to our problem of civil unrest event forecasting from all other tweets and hence reduces the traffic and extra computation of irrelevant tweets. Figure 9 shows the trend of number of tweets being posted in every hour for 7 days sliding window time frame. Figure 9 also reveals that relatively a very small fraction of these tweets are C&C and P&M that increase the significance of filtering these tweets. In C&C

[13] https://web-001.ecs.soton.ac.uk/wo/dataset.

tweets, we use a lexicon based approach and look for the occurrence of various predefined key-terms particularly hashtags (collected by manual inspection and bootstrapping of tweets, related news and articles) relevant to the event. To classify P&M tweets, we train a machine learning classifiers that uses named entities (temporal, spatial, people, organization) of annotated P&M tweets as training dataset. Another input to the classifier is a lexicon of some key-terms related to planning and mobilization. For example- 'join us', 'spread the word', future tense related words. We use these P&M and C&C tweets to build our event prediction model. Research shows that spatio-temporal expressions are discriminatory features for predicting an event. We further include 'topic' as one more discriminatory feature and named entity. There can be multiple named entities in one tweet therefore we define 3 pairs of location-topic (l_i, to_i), topic-time (to_i, ti_i) and location-time (l_i, ti_i) for all expressions. We develop an adaptive version of Budak et al. [9] and compute the correlations among these named entities. To keep only significant and relevant entities, we discard those expressions that occur less than θ times in the dataset (C&C, M&P) where θ varies for all 3 categories of expressions. To find these significant pairs, we extract all expressions of x entity that are at least θ frequent in the dataset $F(x) > \theta N$ and further extract all those expressions of paired entity y that are ψ frequent for at least one x and vice versa. We select only those pairs of named entities that satisfy the following conditions: $F(x, y) > \lceil \psi F(x) \rceil$ and $F(y, x) > \lceil \varphi F(y) \rceil$. We compute the frequency of each pair in bipartite manner and only look for the pairs that are highly correlated and has no decreasing frequency in 7 days of sliding window. Since, these named entities are categorical attributes, we use Chi-Squared distribution to find correlation between two expressions. We define a pair of expressions to be significantly correlated if their *p-value* <0.05 for their respective χ^2 and degree of freedom.

5 Characterization and Classification of Related Work

In this Section, we present a characterization based study on existing literature in the area of civil unrest related event forecasting (refer to Table 1) and hate & extremism detection (refer to Table 2). We analyze all relevant articles and create a list of facets to understand the current state-pf-the-art. Each facet is further classified into sub-categories and have their properties associated with them. Table 1 shows a list of all dimensions (Language, Features and Genre) and their sub-classes for the existing articles in the domain of civil disobedience prediction. Popular social media website allow users to post their content in various languages. The problem of civil unrest event detection can be a region specific as well as on across the globe. Therefore, translating multi-lingual text into English and extracting information (demographic, named entities- temporal, location, topic, people, organization) from translated text is useful for predicting a country or region specific events. Our analysis reveals that in 90 % of the papers researchers have been using English language text and among those in 60 % the articles, the proposed techniques are capable to address multi-lingual text (English, Dutch, French, Portuguese and Spanish). We also observe that 75 % of the articles on civil unrest event forecasting focuses on events specific to a region or country (e.g. Latin America).

Table 1. Various facets for classifying articles on civil unrest prediction

Facets	Categories	Description
Language	English	Mining only English language content
	Non-English	Text consisting Non-English language content
	Others	Testbed consisting of multi-lingual text
Features	Time	Presence of temporal expressions
	Location	Presence of spatial expressions
	Topic	Targeted or central topic of discussion
	Content	Mining text to extract event related information
	Demographic	Demographic and statistical based information
Genre	Regional	Predicting protests happened in one or specific region or country [Latin America]
	Global	Predicting any civil unrest related event happened worldwide

Table 2 shows a list of various dimensions for annotating and existing scholarly articles in the domain of online radicalization detection. Our investigation of the literature reveals that among 37 articles, in 35 papers the proposed method uses contextual metadata as a discriminatory feature to detect hate promoting content[14]. Similarly, in 16 articles, they use demographic information and external links to find relation between two profiles and locate their communities. Similar to civil unrest event prediction related articles, we classify these articles based on the language being addressed. Based on the meta-analysis we find that among 37 articles, 14 articles are capable to mine the content written in Arabic language and script. Among these 14 articles, methods proposed in 7 articles are capable to analyze other non-English language texts as well (German, French). Based on the research aim of articles, we classify previous researches into five genre of online radicalization: Anti-black communities are a form of racism that promotes white supremacy. Jihad communities promote their belief towards Islamic extremism (Sharia Law), also a part of religion based communities where users post hate speech against some religion and promote their ideologies. Hate and Extremism includes the others categories that are undefined and covers political radicalization.

6 ISI Leading Conferences and Journals

The IEEE Intelligence and Security Informatics (ISI) series of conferences[15] is the flagship conference on ISI which started in the year 2003 and has happened

[14] http://bit.ly/1L3x4zV.
[15] http://ieee-isi.org/.

Table 2. Various facets for annotating articles on extremism detection

Dimension	Categories	Description
Features	Text	contextual metadata (Title of Video, Tweet)
	Link	Links between two users (Friend, Follower)
	Demographic	Other demographic and statistical based metadata
Analysis	Content	Detecting the presence of extremist content on web
	User Profile	Identifying hate promoting users
	Community	Locating hidden communities of extremist users
Language	English	Text containing only English language content
	Arabic	content written in Arabic language including scripts
	Others	Posts consisting of any other Non-English language (excluding L2)
Genre	Anti-black	White communities targeting black people
	Jihad	Promotion of Jihad among followers on network
	Terrorism	Social media activities performed by terrorist groups
	Extremism	Promote hatred among various targeted audience
	Religion	Anti-religious content (example- Anti-Islamic Tweets)

in various countries such as USA, Taiwan, Canada, China and Netherlands. ISI is an annual conference and happens normally in the month of May or June. The European Intelligence and Security Informatics Conference (EISIC)[16] is also a leading international scientific conference bringing together researchers and practitioners working in the area of ISI. EISIC has established itself as a premier European conference on counter-terrorism and criminology. It was started in the year 2011 and has happened in various countries such as Greece, Denmark, Sweden, Netherlands and United Kingdom. The Pacific Asia Workshop on Intelligence and Security Informatics (PAISI)[17] is another prestigious forum on ISI which was started in the year 2006 is normally co-located to the Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) every year. Security Informatics[18] is a high quality and high impact open-access Journal from Springer which publishes peer-reviewed articles on ISI. Studies in Conflict & Terrorism formerly known as Terrorism is a peer reviewed based open-accessed journal started in 1992. SICT is ranked in the International Relations category and in the Political Science category of the 2015 Thomson Reuters, 2015 Journal Citations Report.

7 Conclusions

Open source social media analytics for Intelligence and Security Informatics applications is an area that has attracted the attention of several researchers over

[16] http://www.eisic.eu/.
[17] http://www.business.hku.hk/paisi/.
[18] http://www.security-informatics.com/.

the past decade. In this paper, we discuss two important and major application of open source social media intelligence: online radicalization and civil unrest. We perform a comprehensive analysis of literature survey and we observe that even though there has been a lot of work in the domain of event prediction on social media websites, civil unrest event prediction by mining textual content in social media has recently gained the attention of researchers. We also find that the number of research papers on social media analytics for online radicalization detection are much more (about 7 times) than on civil disobedience detection[19]. Our analysis also reveals that Twitter plays a major role in facilitating mobilization and planning of civil unrest events in comparison to other social media platforms. It is interesting to observe that despite being the most popular video sharing website, YouTube has not been used in any of the existing research for protest planning or prediction. On the contrary, research shows that YouTube is the most widely used platform for online radicalization, hate and extremism promotion. We conduct a meta-analysis of exiting literature in the domain of online radicalization detection and civil unrest even prediction on social media platforms. Our analysis reveals a variety of machine learning, information retrieval and data mining approaches that has been used in past to investigate solution for these problems. Based on our analysis, we find that Clustering, Logistic Regression and Dynamic Query Expansion are the most commonly used techniques to predict upcoming events related to civil unrest or protest. Named Entity Recognition (NER) is a most common and widely used component in all above event prediction models. Graph modeling and Ensemble Learning are also some techniques adopted by researchers for the problem of event forecasting. Text Classification KNN, Naive Bayes, SVM, Rule Based Classifier, Decision Tree, Clustering (Blog Spider), Exploratory Data Analysis (EDA) and Keyword Based Flagging (KBF) are the most commonly used techniques to identify extremism and hate promoting content on Internet. However, for navigating through links and locating users with similar interest, Topical Crawler and Link Analysis (Breadth First Search, Depth First Search, Best First Search) are common graph traversal algorithms in use. Language modeling, character level n-gram and Boosting are other techniques used for mining textual data and classifying hate promoting content on web (social media, blogs and forums).

In this paper, we also present two case studies as two applications of open source social media analytics for Intelligence and Security Informatics. We present a classification framework for the problem of identifying hate promoting content on YouTube and a topical crawler based approach to locate such extremist users and their hidden communities sharing a common agenda. We conduct a series of experiments on YouTube videos and channels by varying various algorithmic parameters such as the similarity threshold for the language modelling based text classifier and n-grams. We conclude that by performing social network analysis on network graphs, we are able to locate hidden communities. We also identify the users who play major roles in the communities and

[19] http://bit.ly/1L3x4zV.

have highest centrality among all. We reveal the communities by dividing the network graph into clusters formed by similar users (refer to our previous work [2,4] for more details). For civil disobedience forecasting, we propose an approach for early detection of these events. In our approach we present a case study on immigration related event- Christmas Island Hunger Strike. To investigate the effectiveness of our approach, we conduct experiments on real world dataset downloaded from Southampton WebObservatory. We train a multi-class classifier to filter event related tweets (crowd-buzz and mobilization) which certainly improves the accuracy the accuracy of prediction and reduces the computational cost. We develop a frequency based model on these event related tweets (semantically rich) and find those pairs of named entities (spatial location, topics and temporal) expressions that are significantly correlated. We perform the χ^2 and p-value distribution on these pairs of named entities expressions and conclude that by detecting trend analysis of spatial, temporal and topic based entities in sliding window (7 days), we can predict civil unrest related events with high confidence score. We further perform an in-depth meta-analysis of existing literature and perform a characterization based classification of articles. Based on our analysis we find that contextual based metadata (title, description, comments) is most commonly used feature for identifying hate promoting content. However, to find the relation between user channels and to locate their hidden communities, demographic information and activity feeds of user profile are common discriminatory features. We observe that many of the existing techniques are capable to mine multi-lingual text such as Arabic and capture relevant information.

References

1. Agrawal, S., Sureka, A.: Copyright infringement detection of music videos on YouTube by mining video and uploader meta-data. In: Bhatnagar, V., Srinivasa, S. (eds.) BDA 2013. LNCS, vol. 8302, pp. 48–67. Springer, Heidelberg (2013)
2. Agarwal, S., Sureka, A.: A focused crawler for mining hate and extremism promoting videos on YouTube. In: 25th ACM Conference on Hypertext and Social Media (HT), pp. 294–296 (2014)
3. Agarwal, S., Sureka, A.: Learning to classify hate and extremism promoting tweets. In: Intelligence and Security Informatics Conference (JISIC), pp. 320–320 (2014)
4. Agarwal, S., Sureka, A.: Topic-specific YouTube crawling to detect online radicalization. In: Chu, W., Kikuchi, S., Bhalla, S. (eds.) DNIS 2015. LNCS, vol. 8999, pp. 133–151. Springer, Heidelberg (2015)
5. Agarwal, S., Sureka, A.: A topical crawler for uncovering hidden communities of extremist micro-bloggers on tumblr. In: 5th Workshop on Making Sense of Microposts (MICROPOSTS) (2015)
6. Agarwal, S., Sureka, A.: Using common-sense knowledge-base for detecting word obfuscation in adversarial communication. In: Workshop on Future Information Security (FIS) (2015)
7. Agarwal, S., Sureka, A.: Using KNN and SVM based one-class classifier for detecting online radicalization on Twitter. In: Natarajan, R., Barua, G., Patra, M.R. (eds.) ICDCIT 2015. LNCS, vol. 8956, pp. 431–442. Springer, Heidelberg (2015)

8. Aggarwal, N., Agarwal, S., Sureka, A.: Mining YouTube metadata for detecting privacy invading harassment and misdemeanor videos. In: Privacy, Security and Trust (PST), pp. 84–93 (2014)
9. Budak, C., Georgiou, T., Agrawal, D., El Abbadi, A.: Geoscope: online detection of geo-correlated information trends in social networks. Proc. VLDB Endow. **7**, 229–240 (2013)
10. Compton, R., Lee, C.: Detecting future social unrest in unprocessed Twitter data: emerging phenomena and big data. In: Intelligence and Security Informatics (ISI), pp. 56–60 (2013)
11. Fu, T., Huang, C.N., Chen, H.: Identification of extremist videos in online video sharing sites. In: 2009 IEEE International Conference on Intelligence and Security Informatics, ISI 2009, pp. 179–181, June 2009
12. Hua, T., Lu, C.T., Ramakrishnan, N.: Analyzing civil unrest through social media. Computer **46**(12), 80–84 (2013)
13. Kwok, I., Wang, Y.: Locate the hate: detecting Tweets against blacks. In: AAAI (2013)
14. Qazvinian, V., Rosengren, E., Radev, D.R., Mei, Q.: Rumor has it: identifying misinformation in microblogs. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, Stroudsburg, PA, USA, pp. 1589–1599 (2011)
15. Ramakrishnan, N., Butler, P., Muthiah, S.: 'Beating the news' with embers: forecasting civil unrest using open source indicators. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2014, pp. 1799–1808. ACM, New York (2014)
16. Wang, M., Alan, C.G.: Intelligence and security informatics. In: Pacific Asia Workshop (PAISI) (2011)