LITERATURE REVIEW (PHD COMPREHENSIVE EXAMINATION)

Applying Social Media Intelligence for Predicting and Identifying On-line Radicalization and Civil Unrest Oriented Threats

Swati Agarwal
PhD Adviser: Dr. Ashish Sureka

Abstract
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 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 metanalysis of about 100 conference and journal papers to gain a better understanding of existing literature.

Keywords: Intelligence and Security Informatics; Machine Learning; Mining User Generated Content; Event Forecasting; Hate and Extremism; Law Enforcement Agency; Online Radicalization; Social Media Analytics

1 Introduction
Over the past decade, social media has emerged into a dynamic form of world-wide interpersonal communication. It facilitates users for constant and continuous information sharing, making connections and conveying their thoughts across the world via different mediums. For example- social networking (Facebook), micro-blogging (Twitter, Tumblr), image sharing (Imgur, Flickr) and video hosting and sharing (YouTube, Dailymotion, Vimeo). Due to high reachability and popularity of social media websites worldwide, organizations use these websites for planning and mobilizing events for protests and public demonstrations [1]. The study of civil unrest reveals that now most of the protests are planned and mobilized in much advance [2] [3]. Crowd-buzz about these protests over social media is a rich source for civil unrest forecasting. Traditionally, newspapers have been used a primary sources for such analysis and prediction. However, the speed and flexibility of publication on social media platforms gained the attention of various organizations for planning and making announcements of various protests, strikes, public demonstrations and riots. Similarly, simplicity of navigation, low barriers to publication (users only need to have a valid account on website) and anonymity (liberty to upload any content without revealing their real identity) have led users to misuse these websites in several ways by uploading offensive and illegal data[1]. Popular social media websites, blogs, forums are frequently being misused by many hate groups to promote online radicalization (also referred as cyber-extremism, cyber-crime and cyber hate propaganda). Research shows that extremist groups put forth hateful speech, offensive and violent comments and messages focusing their mission. Many hate promoting groups use popular social media websites to promote their ideology by spreading extremist

[1]http://www.policechiefmagazine.org/magazine/index.cfm
content among their viewers [4] [5]. They communicate with other such existing groups and form their virtual communities on social media sharing a common agenda. They use social networking as a medium to facilitate recruitment of new members in their group by gradually reaching wide audiences that help to persuade others to violence and terrorism [6]. Researchers from various disciplines like psychology, social science and computer science have constantly been developing tools and proposing techniques to counter and combat these problems of online radicalization and generating early warning for civil unrest related events.

Presence of extremist content and planning of civil unrest are major concerns for the government and law enforcement agencies [7]. Online radicalization has a major impact on society that contributes to the crime against humanity and mainstream 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. Monitoring the presence of such content on social media and keeping a track of this information in real time is important for security analysts who work for law enforcement agencies. Figure 2 illustrates top 3 Twitter accounts of hate promoting users posting extremist content very frequently and having a very large number of followers [8]. 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.

Over past 10 years, since 2005 several approaches, techniques, algorithms and tools have been proposed to mine the data and bring solutions to the problems which are encountered by Intelligence and Security Informatics (ISI)[9] [10]. The aim of the study presented in this paper is to conduct a systematic literature survey of those previous techniques as documented in scholarly articles. Our goal is to do a comprehensive analysis of these articles to better understand the state-of-the-art, research gaps, techniques and future directions.

1.1 Technical Challenges

The problem of automatic identification of online radicalization (hate promoting content, users and hidden communities) and prediction of civil unrest related events (protests, riots, public demonstrations) is an important problem for ISI researchers. However, due to the dynamic nature of social media platforms, identifying such content, locating users and predicting events by keyword-based search is overwhelmingly impractical. The volume of content being posted on social media platforms makes it challenging for security analysts to discover such content manually. For example, over 6 billion hours of video are watched each month on YouTube. 100 millions of people perform social activities every week and millions of new subscriptions...
are made every day. According to Tumblr statistics 2015 [3], over 219 million blogs are registered on Tumblr (second most popular micro-blogging website) and 420 million are the active users. 80 million posts are being published everyday, while the number of new blogs and subscriptions are 0.1 million and 45 thousands respectively [11]. Presence of large volume and multi-media content makes it difficult to find common patterns and trends in data. Security analysts also encounters problems in visualizing the data and relationship among users uploading such content. Manual inspection and keyword based flagging increase the number of false positive and reduce the efficiency of proposed approaches.

Further, textual posts on social media websites are user generated content which is unstructured and informal. User generated data contains noisy content such as incorrect grammar, misspell words, internet slangs, abbreviations and text containing multi-lingual script. Presence of low quality content in contextual metadata increases the complexity of problem and poses technical challenges to text mining and linguistic analysis [12] [13] [14]. Figure 1 shows an example of a Tumblr post that contains multi-lingual text (English and Arabic).

Adversarial behavior on social media also poses a technical challenge to filter the content related to hate promoting and event planning for civil disobedience. Despite providing the feasibility of posting any kind of content popular social media websites have several guidelines for uploaders posting content on the website. According to those guidelines users are not allowed to post any illegal or unethical content on the website. Due to these reasons users post information which might seem genuine but leads to sensitive information. For example, title and description of a video does not contain any suspicious term while the video has the content related to hate promotion. Similarly, a tweet might have no term related to event planning but the URL present in the tweet redirects to an external page containing information related to a public demonstration.

1.2 Our Contributions

This paper falls under the aims and scope of Intelligence and Security Informatics (ISI). We have conducted an in-depth and rigorous literature survey on a sub-topic within ISI. Our literature review is comprehensive and provides insights useful to the ISI research community. To the best of our knowledge this paper is the first such survey of existing literature on social media in the domain of automated techniques for Online Radicalization detection and generating early Warning or predicting civil unrest related events. We propose a one class classification architecture across several dimensions (social media platforms being used as a data source, countering issues of online radicalization and civil unrest, machine learning techniques) to classify scholarly articles that fall under the scope of focus of our research (refer to section 2). We perform an in-depth characterization and classification based upon meta analysis of existing researches. We analyze every article and inspect commonly used techniques, identify trends and find research gaps.

Road Map of Literature Survey

The literature review presented in this paper is divided into multiple sections. The paper is organized as follows: Section 2 discusses the scope and focus of our research problem where we define the topics and subtopics covered in the literature review. Section 3 describes the general framework and the process followed to collect relevant scholarly articles for conducting literature survey. In Section 4, we discuss the analysis of number of publications over a time period, type of paper (full, short, poster paper and journal) and their associated venues (conference, journal). In Section 5, we perform an in-depth characterization and classification of articles based upon meta analysis (machine learning techniques, experimental data source). We also present some statistics of existing literature in various dimensions such as discriminatory features used for classification and languages being addressed in solution approach followed by the concluding remarks in Section 6.

2 Research Scope and Focus

Figure 3 defines the topic and scope of our literature review. As shown in Figure 3, our research focus is at the intersection of the following three fields:

1. Online Social Media Platforms
2. Intelligence and Security Informatics
3. Text Mining and Analytics

We restrict our analysis to studies on mining user generated content on social media platforms like Twitter (micro-blogging website), YouTube (video-sharing website), blogs and discussion forums and not on documents or intranet sites within an enterprise or organization. Our focus is on mining information which is publicly available (open-source intelligence). Web 2.0 and Social Media platforms contain content belonging to various modalities like image, audio, video and text. We restrict our analysis on the application of machine learning and information retrieval techniques on free-form textual data and not sound, image and video data present in social media platforms. Intelligence and

[3] https://www.tumblr.com/about
security informatics is a vast field and for the literature survey presented in this paper, we focus on two important applications: online radicalization and civil unrest.

3 General Research Framework For Literature Review

Figure 4 illustrates the sequence of steps followed by us to collect papers for conducting our literature review. As shown in Figure 4, we begin by creating a list of key-terms representing our topic or problem area. For example, some of the search key-terms to retrieve relevant papers are: 'civil unrest/protest', 'event forecasting', 'early warning', 'extremism detection', 'online radicalizing community detection'. We search relevant articles using Google Scholar[4] which is a well-known and widely used web-based search engine for finding scholarly literature. We use Google Scholar as it has a good coverage and index of articles and also provides a powerful mechanism to explore related work, citations and authors. Google Scholar also provides data on how often and how recently a paper has been cited which is also a useful metric (judging the impact) for conducting our literature survey. We go through the Title, Keywords and Abstract of the articles retrieved as a result of key-term based search as well as through related article and cited by links provided for each article in the search result. We perform a meta analysis on article and determine the relevance of the article to our topic or focus area. We perform a one class classification (Rule Based Classifier) on each article and check if it meets the scope and focus defined in Section 2. During classification, we typically mine the Title, Keywords and Abstract and accordingly save the article in our database of collection of closely related work.

4 Survey of Number of Publications and Venue

We limit our literature review to papers published in peer-reviewed conference and workshop proceedings and journals (and indexed in bibliographic databases such as IEEE Xplore, ACM Digital Library and Springer). We do not include granted patents or patent applications, books, lecture notes or slides, software products and information websites. We record the year, venue (conference or journal name) and type of every paper (full, short or poster paper in-case of a conference). As shown in Figures 5 and 6, we classify each paper (civil unrest and online radicalization) in terms of its type (y-axis), year (x-axis) and the venue

[4]https://scholar.google.co.in/
Figure 4: General Research Framework For Literature Review Process Followed in our Survey

Figure 5: A Heatmap Representation of the Variance in Number of Publications and Conference Venues [Civil Unrest Related Event Forecasting] Over a Time Period

(value in the cell). The graphs in Figures 5 and 6, also shows a gradient (gradual color change) in which each cell is background shaded with a color. The color in the cell represents the number of papers published with the period and paper type represented by the cell. Usage of color ramp or progression as one dimension helps us in understanding the trends (growing, stable, and declining) in paper publications in our focus area.

4.1 Event Forecasting or Early Warning for Civil Unrest Related Events

Figure 5 reveals that there are a total of 8 publications in both conferences and journals in the area of civil unrest forecasting by mining social media content. Figure 5 also reveals that maximum number of papers are published in 2014 including one journal [15], 3 full papers [2] [16] [17] and one short paper [3]. During the time period of 2010 to 2013, we only find 2 pub-
| Journal                               | Full Paper                                                                 | Short Paper                                                                 | Poster                                                                 |
|---------------------------------------|----------------------------------------------------------------------------|----------------------------------------------------------------------------|------------------------------------------------------------------------|
| Communication Law & Policy            |                                                                           |                                                                           |                                                                        |
| Journal of Criminal Justice and Popular Culture | Confronting right wing extremism and terrorism in the USA                      |                                                                           |                                                                        |
| 1. Intelligent Systems 2. Computer Mediated Communication | Intelligence and Security Informatics (ISI)                                     |                                                                           |                                                                        |
| Security Informatics (SI)             | 1. PACIS 2. ICIS 3. CLC 4. Homeland 5. Public Choice 6. BSD Review         |                                                                           |                                                                        |
| 1. Mobilization: An International Quarterly 2. IJHCS | New York State Communication Association                                      |                                                                           |                                                                        |
| 1. ISI 2. Dynamics of Asymmetric Conflict 3. ECSS |                                                                                |                                                                           |                                                                        |
| 1. CLAWS 2. MME 3. HJC 4. NMS 5. Observer India | 1. ISI 2. ASONAM 3. Communities and technologies                              | 1. ISI 2. Intelligent Agent & Multi-Agent Systems                       |
| 1. CM 2. CDS 3. JASIST 4. TAVPA      | 1. Information Retrieval Technology 2. ASONAM                               |                                                                           |                                                                        |
| 1. Studies in Conflict & Terrorism 2. News Media & Society | EISIC                                                                       |                                                                           |                                                                        |
| 1. SI 2. DKE 3. Asia Pacific 4. Journal of Research | 1. ISI 2. EIDWT                                                            | Homeland Security                                                       |                                                                        |
| 1. AAAI 2. arXiv 3. NCC 4. WebSci 5. USMA 6. ISS 7. ACT |                                                                                |                                                                           |                                                                        |
| 1. Online Journal of Communication and Media Technologies 2. SI 3. TAPV | Cyberterrorism                                                              | 1. JISIC 2. HyperText                                                    | Making Sense of Microposts                                             |
| 1. DNIS 2. ICDCIT                     |                                                                           |                                                                           |                                                                        |

Figure 6: A Heatmap Representation of the Variance in Number of Publications and Conference Venues [Online Radicalization] Over a Time Period
lications one every year while there is no publication in 2011 and 2012 [18] [19]. We observe that over past 6 years, no conference or journal have more than one publication on this research problem except SIGKDD (has 2 full papers in 2014) [2] [17]. By Figure 5, we conclude that civil disobedience event forecasting problem has recently gained the attention of researchers.

4.2 Identification of Online Extremist Content, Users and Communities

Similar to Figure 5, Figure 6 shows the number of publications and venue for existing literature conducted in the domain of online radicalization detection. Figure 6 reveals that there has been a lot of work in the domain of hate and extremism detection on social media. There are many journal papers and conference proceedings available over past 15 years. We observe that this problem is not very new. However, this problem is important and not fully solved. Researchers have constantly been working on this problem and publishing papers in reputed conferences and journals. Here, the first few publications during the time period of 2001 to 2005, are not using social media as a platform for conducting experiments and have been published in crime and law specific venues (both international and national) [20] [21]. For example, Confronting right wing extremism and terrorism in the USA (conference 2003) is a national conference exclusively for researches in the domain of hate and terrorism detection [22]. Figure 6 also reveals that papers in online radicalization detection domain are published in wide range of conferences (32) and journals (23). Maximum number of papers are published in ISI conference[5] (more than 6) and Security Informatics journal[6] (more than 4) which are two of the most core venues for the research domain and problem [23] [24] [25] [26] [27] [28]. Figure 6 shows that the maximum number of journal and total (including full and short papers) publications are in 2009 [26] [29] [30] [31]. While, maximum number of conference regular papers are published in 2013 (7 full papers) [32] [33] [34] [35] [36] [37] [38] and 2006 (6 full papers) [25] [39] [40] [41] [42]. There are only 3 poster papers published in 2014 (JISIC [6] and HT [4]) and 2015 (Making Sense of Microposts [11]). Heatmap shown in Figure 6 also reveals that there is no publication in 2004.

5 Characterization and Classification of Articles Based Upon Meta Analysis

In this part of literature review, we present a characterization based study on previous researches done in the area of civil unrest related event forecasting (refer to Table 2) and online radicalization detection (refer to Table 4). We analyze each paper and create a list of all dimensions to demonstrate the statistics. We also present statistics of existing literature that use social media platforms as data source for conducting experiments. Each dimension is further classified in subcategories and have their properties associated with them. These dimensions are as follows:

1. Data Source
2. Techniques
3. Features
4. Evaluation
5. Type of Analysis
6. Language
7. Genre
8. Region (only for online radicalization)

5.1 Social Media Websites as Data Source

As discussed in Section 4, there are only 8 existing publications in the domain of civil disobedience event
Table 1: Statistics of Number of Publications Using Various Social Media Platforms as Data Source for Conducting Experiments over Past Decade

| Data Source       | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 |
|-------------------|------|------|------|------|------|------|------|------|------|------|
| Twitter           | 0    | 0    | 0    | 1    | 0    | 0    | 6    | 7    | 1    | 1    |
| YouTube           | 0    | 0    | 2    | 3    | 2    | 2    | 2    | 1    | 1    |      |
| Facebook          | 0    | 0    | 0    | 1    | 0    | 0    | 3    | 2    | 0    | 0    |
| Tumblr            | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 1    |      |
| Blogs             | 2    | 2    | 1    | 1    | 0    | 0    | 0    | 0    | 0    |      |
| Forums            | 2    | 0    | 0    | 2    | 1    | 1    | 0    | 1    | 1    | 0    |
| Webpages          | 2    | 0    | 0    | 1    | 0    | 0    | 0    | 0    | 0    |      |
| News Articles     | 1    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0    |      |
| Image Hosting     | 0    | 0    | 0    | 0    | 0    | 0    | 1    | 0    | 0    |      |
| Gaming/Virtual World | 0   | 0    | 1    | 0    | 0    | 0    | 0    | 0    | 0    |      |

Figure 8: Number of Publications Over a Period of Time using One or More Social Media Platforms For Online Radicalization Detection

Forecasting and more than 40 publications in the area of online extremism detection. Therefore, for the problem of event forecasting for civil unrest related events, we take publication IDs on x-axis and data source on y-axis; colors represent the different social media platform and other data sources. Figure 7 reveals that Twitter and News media reports are the most widely used data sources for event forecasting [1] [2] [18]. Apart from the news media articles, micro-blogging websites are a rich source of information. In Figure 7, we observe that the papers using a single source of data are using micro-blogging websites (Twitter and Tumblr) [3] [16] [18]. Most of the researchers have used multiple data sources (blogs, forums, news media and other social networking websites) for information extraction [1] [2]. We also observe that despite being most popular video hosting and sharing website and widely used YouTube has not been used in any of the existing research for protest planning or prediction.

For online radicalization detection, we identify the number of publications (y-axis) using various data sources (legends) and experimental dataset over past decade (x-axis). Figure 8 reveals that there has been some work on all popular social networking platforms, micro-blogging websites [3] [16] [18], video sharing [30] [43] [44] [45] and image hosting websites [46]. Many researchers have used blogs [39] [47], forums [26] [48], web documents [48], news articles [49] and online gaming websites [50] for hate and extremism detection (refer to Table 1 for statistics). Figure 8 reveals that initially in 2006, most of the researches were conducted using blogs, discussion forums, news articles and extremist websites as a data source. YouTube, Facebook and Twitter were founded in 2004, 2005 and 2006 respectively which gradually became popular among extremist users for posting hate promoting content. Previous research shows that YouTube is most widely used platform for hate promotion. Figure 8 also reveals YouTube is being used as a rich source of extremist data for online radicalization detection over past 7 years [30] [43] [44] [45] and image hosting websites [46]. Micro-blogging websites are very popular and widely used platforms among its users (including terrorist organization and hate promoting groups) for conveying messages and forming communities. Figure 8 shows that recently, Twitter has gained the attention of researchers and we find many publications in past 3 years that are using Twitter for identifying ex-
tremism content, users and communities [6] [35] [51] [52]. We also observe that despite being second most popular micro-blogging website, we find only publication that is using Tumblr (founded in 2007) as a data source for extremism content identification [11]. Table 1 shows that there are 15 papers that have used YouTube as a data source and similarly 16 papers using Twitter, 8 papers using discussion forums and 6 papers using blogs for collecting experimental dataset to counter and combat online radicalization.

5.2 Machine Learning and Data Mining Techniques

Figures 9 and 10 illustrate the machine learning and information retrieval techniques used in previous papers of civil unrest event prediction and online radicalization respectively.

Similar to Figure 7, Figure 9 illustrates the proposed methods and techniques used in each paper individually. Ramakrishnan [2] proposed five different techniques for event forecasting for five different kinds of data and models (volume based, opinion based, tracking of activities, distribution of events and cause of protest). Unlike Ramakrishnan [2], in Figure 9, we find that most of the researchers have applied ensemble learning on multiple data mining and machine learning techniques to achieve better accuracy in prediction [1] [16] [19].

Clustering, Logistic Regression and Dynamic Query Expansion are the commonly used techniques to predict upcoming events related to civil unrest or protest. We go through the methodology Section of each paper and observe that named entity recognition is a common phase for all approaches illustrated in Figure 9. In named entity recognition, they extract several entities present in the contextual metadata (tweets, Facebook comment, News article) [53]. These entities could be spatiotemporal expressions, topic being discussed in posts (refer to Table 2 for full list of entities and features). List of all entities and keywords is dynamically expanded by using Dynamic Query Expansion method where they find similar and relevant keywords/entities using external lexical source (for example- WordNet[7], VerbOcean[8]). Dynamic Query Expansion is an iterative process and converges once the keywords are stable. They further perform several clustering and classification techniques on these entities and text to predict upcoming events.

Another popular technique used for event forecasting is graph modeling. Feng Chen et.al. [17] uses a heterogeneous graph modeling as keyword enrichment and pre-processing. To achieve accurate results in event forecasting they use only filtered entities for NPHGS approach. According to Feng Chen et.al. [17], a heterogeneous graph is defined as a network consisting of nodes, edges and relations where nodes are the entities extracted using named entity recognition. There can be multiple types of nodes equivalent to number of entities extracted (topic, temporal, spatial, organization etc). Edges are the link between two entities and relation defines the feature vector between two entities. For example, one relation between a Twitter user $U$ (entity: people) and a term $T$ (entity: topic) can be number of tweets by user $U$ on topic $T$. Entities having high relevance and polarity score above certain threshold are filtered and used in next phase of NPHGS.

Unlike Figure 9, due to large number of papers in online radicalization domain, we present machine learning technique over a timeline instead of individual presentation for each paper. Figure 10 reveals that text

---

[7] https://wordnet.princeton.edu
[8] http://demo.patrickpantel.com/demos/verbocean/
classification KNN, Naive Bayes, Support Vector Machine, Rule Based Classifier, Decision Tree, Clustering (Blog Spider), Exploratory Data Analysis (EDA), Topical Crawler/Link Analysis (Breadth First Search, Depth First Search, Best First Search) and Keyword Based Flagging (KBF) are the most widely used techniques for online radicalization detection on social media websites [30] [40] [51] [52]. For all techniques, there are different sets of discriminatory features which we will discuss in detail in Section 5.3. Social networking websites, micro-blogging websites and video sharing websites are amongst the largest repositories of user generated content on web. Therefore, Text classification (automatic and semi-supervised learning), clustering (unsupervised learning), EDA and KBF approaches are very well known techniques and commonly used for identifying extremist content on social media [37] [54]. While Topical crawler and Link analysis are the techniques used for crawling through navigation links and identifying similar users and locating hidden communities on social media websites [11] [48] [55]. The topical crawler is a recursive process that adds and removes nodes after each iteration. It starts from a seed node, traverses in a graph navigating through some links and returns all relevant nodes to a given topic. Breadth First Search, Depth First Search and Best First Search are different ways to pick neighbors and navigate through external links. These links and neighbors are different for different social networking websites. For example, if user $u$ posted a tweet $t$ then a neighbor can be a follower of user $u$, or users liking and re-tweeting or re-blogging (in case of Tumblr) that post. Similarly in YouTube, if user $u$ posted a video $v$ then a link can re-direct to a user channel who subscribed $u$ or posted a comment on video $v$.

Language modeling, n-gram, Boosting are other techniques for classifying textual data as hate promoting based upon several discriminatory features [4]. TREC is a document ranking algorithm to search and rank documents available at Text REtrieval Conference. Bermingham et. al. [29] uses TREC algorithm to perform sentiment analysis of comments posted on extremist videos on YouTube. However, they find that due to the difference in nature of data, TREC is not useful to detect subjectivity in YouTube comments and makes the results unreliable. Since comments posted on YouTube are opinion based while the documents in TREC corpus are based upon facts.

**OSLOM (Order Statistics Local Optimization Method)** is a clustering algorithm designed for graphs and networks which has been used by some researchers to locate groups and communities of extremist users sharing a common agenda. OSLOM is an open source visualization tool capable to detect hidden communities in a network accounting for edge directions, edge weights, overlapping communities, hierarchies and community dynamics [34] [35]. OSLOM locally optimize the clusters and consists of following three phases:

1. Iterative process to identify significant clusters of nodes

**Figure 10: Machine Learning Techniques Used By Researchers in Existing Literature of Online Radicalization and Hate Promotion Detection**
Table 2: List of Various Dimensions and their Associated Properties Used for Annotating Literature Articles Conducted in the Domain of Civil Unrest Related Event Forecasting

| Dimension Categories | Description |
|----------------------|-------------|
| **Features**         |             |
| F1 Temporal          | Presence of time related expressions |
| F2 Spatial           | Presence of location based expressions |
| F3 Topic             | Presence of targeted topic or domain related expressions |
| F4 Content           | Mining text to extract event related information |
| F5 Demographic       | Other demographic and statistical based metadata |
| **Evaluation**       |             |
| E1 K-Cross Validation| Optimizing the output of forecasting by splitting data into k-samples |
| E2 Precision         | Evaluates the exactness of forecasting results |
| E3 Recall            | Evaluates the completeness of results |
| E4 NA                | No evaluation technique is mentioned |
| **Analysis**         |             |
| A1 Content           | Mining only textual content for feature extraction and event forecasting |
| A2 Community         | Mining user profiles and networks for extracting event related information |
| **Language**         |             |
| L1 English           | Conducting experiments only on English Language 1 tweets |
| L2 Non-English       | Posts consisting of any Non-English language content |
| L3 Others            | Testbed consisting of multiple languages content |
| **Genre**            |             |
| G1 Protest-Country Specific | Conducting Experiments for the protests happened in one or specific countries/region [Latin America] |
| G2 Global            | Performing Event forecasting on any civil unrest related event happened worldwide |

Table 3: Summary of Existing Literature in the Domain of Civil Unrest Related Event Forecasting

| Authors            | Year | Genre | Analysis | Evaluation | Language | Features |
|--------------------|------|-------|----------|------------|----------|----------|
| Colbaugh et. al.   | 2010 | G1 G2 | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | F1 F2 F3 F4 F5 |
| Hua et. al.        | 2013 | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ |
| Ramakrishnan et. al. | 2014 | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ |
| Jiejun et. al.     | 2014 | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ |
| Chen et. al.       | 2014 | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ |
| Compton et. al.    | 2014 | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ |
| Filchenkov et. al. | 2014 | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ |
| Muthiah et. al.    | 2015 | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ | ✓ ✓ ✓ ✓ ✓ |

2 Analyzing resultant clusters and identification of internal structure and overlapping clusters (if possible)

3 Identification of hierarchical structure of clusters.

5.3 Characterization and Classification of Articles- Early Warning and Prediction of Civil Unrest Related Events

Table 2 illustrates various facets and their associated properties used by us for performing an in-depth characterization of existing literature. Table 3 shows the summary of this characterization performed on all 8 papers. Features mentioned in the Table 2 are mainly of three types: entities, content analysis and demographic information. Entities are extracted using named entity recognition (refer to Section 5.2) and are pre-defined for domain specific problem. These entities are temporal expressions (today, tomorrow, 10pm), spatial location expressions (USA, in front of white house, geocodes) and the topic being discussed in posts (migration, protest). Content analysis includes the mining and analyzing contextual metadata. For example, presence of hashtags, @user mentioned in a tweet. Demographic information are the statistical based measures. For example- number of tweets posted by a user, age and location of user. Table 3 reveals that 75% of papers have used spatiotemporal features as discriminatory features for predicting events. While rest of the 25% uses only timeline as a discriminatory features and uses event tracking for prediction. Table 3 also reveals that all the papers using spatiotemporal features also use content analysis to extract more relevant information from raw text. Content analysis feature includes searching for domain specific key terms. Precision, Recall and K-cross validation are standard information retrieval techniques to evaluate the performance of proposed techniques. Table 2 reveals the common evaluation measures used by researchers. Precision computes the ‘exactness’ of proposed method while recall computes the ‘completeness’ of solution approach. K-cross validation splits the testbed into K parts and evaluates the model iteratively for each part. K-cross validation is used to optimize the results. Table 3 reveals that 65% of the studies use Precision measure to evaluate the performance of their proposed approach among which 25% papers use both recall and precision measures. While in 2 papers no evaluation method is defined and in only 1 paper, authors use k-cross validation method to measure the effectiveness of their approach.
Table 4: List of Various Dimensions and their Associated Properties Used for Annotating Literature Articles Conducted in the Domain of Online Radicalization Detection

| Dimension | Categories | Description |
|-----------|------------|-------------|
| **Features** | F1 Text | Textual based content of the posts (Title of Video, Tweet, Facebook Comments etc.) |
| | F2 Link | Links between two user profile (Subscription in YouTube, Follower in Twitter and Tumblr etc.) |
| | F3 Demographic | Other demographic and statistical based metadata |
| **Evaluation** | E1 Precision | Evaluates the exactness of forecasting results |
| | E2 Recall | Evaluates the completeness of results |
| | E3 F-Score | Weighted harmonic mean of Precision (E1) and Recall (E2) |
| | E4 Accuracy | Evaluates the correctness of the technique |
| | E5 K-cross Validation | Optimizing the output of forecasting by splitting data into k-samples |
| | E6 SNA | Performs Social Network Analysis to show the findings in community detection |
| | E7 User Based | Evaluation performed by external users |
| | E8 NA | No evaluation method is defined. |
| **Analysis** | A1 Content | Mining only textual content for feature extraction and event forecasting |
| | A2 User Profile | Mining user profiles metadata for extracting event related information |
| | A3 Community | Mining linked profiles and their communities |
| **Language** | L1 English | Conducting experiments only on English Language Tweets |
| | L2 Arabic | Testbed consisting of content written in Arabic language (scripted or language text) |
| | L3 Other Non-English | Posts consisting of any other Non-English language (excluding L2) content (German, French etc.) |
| **Region** | R1 US Domestic | Content posted targeting US issues or radicalization originated from US domestic regions |
| | R2 International | Researchers focusing on global or international radicalization or extremism |
| | R3 Others | Radicalization originating or targeting other groups and regions worldwide (example- Middle Eastern, Latin America) |
| **Genre** | G1 Anti-black | Content posted by white supremacy communities targeting black people |
| | G2 Jihad | Groups posting content for promoting Jihad among their viewers |
| | G3 Terrorism | Content posted by terrorism group or social networking activity performed by terrorists |
| | G4 Hate and Extremism | Content posted in order to promote hate and extremism among various targeted audience |
| | G5 Religion | Content posted against a religion (example- Anti-Islamic Tweets) |

Right prediction of an event depends upon the content present in the documents, tweets and other contextual metadata. However, we observe that some researchers also use profile information of posters and groups of these posters talking about an event for event forecasting. Therefore, we categorize these papers based upon two classes of analysis: content and community. Content, where author mine only contextual metadata for feature extraction and event prediction. Community, where authors mine both user profiles and network (links) for extracting event related information. Table 3 reveals all papers have used content analysis for predicting events related to protest and civil disobedience. While in 50% of the articles, authors use profile information of users posting event related content and a bunch of users talking about the event. This information includes information about geographical location (geocodes), frequent '@' mentioned (to find other connected user profiles), frequent hashtags used (to expand the list of domain and event specific keywords).

Popular social media website allow users to post content in multiple language being spoken across the world. Translating multi-lingual text into English language text and extracting information from it is useful when predicting a country or region specific events. We classify these papers into 3 categories based upon the language being addressed for event prediction. Table 3 reveals that in 90% of the papers authors have worked on English language text among which in 60% the papers they have addressed multiple languages (both English and Non-English) text. Dutch, French, Spanish and Portuguese are the languages that were addressed in maximum papers. We also find one paper where authors have worked on only Spanish language tweets.

We classify articles into different genre based upon the targeted region defined in articles. Based upon our analysis we define two genre for region based classification: Country Specific (example- Latin America, Africa) and Global. Table 3 reveals that 75% of papers focus on events specific to a region or a country among which only one article addressing the events happening across all local and global regions.

5.4 Characterization and Classification of Articles—Identification of Extremist Content, Users and Hidden Communities

Table 4 shows various dimensions and their associated categories used by us to classify existing scholarly articles in the domain of online radicalization detection.

In online radicalization detection literature review, we collect articles focusing towards identification of extremism content and locating hate promoting users and communities. Therefore we classify these articles based upon three discriminatory features being used by researchers. Text defines the contextual metadata
of a post. For example, content of a tweet, title of a video, tags present in the description of a video etc. Link defines the relation between two user profile which can be different for different social media platforms. For example, "post re-blogged by" in Tumblr, a "follower" in Twitter, "video liked by" in YouTube etc. Demographic information are the statistical based metadata of post and user profiles. For example, number of comments on a video, number of favorites on a tweets, notes on a Tumblr post etc. Figure 12 shows the statistics of these features used by various articles collected in the domain of online radicalization. Figure 12 reveals that the number of papers using text, link and demographic information are 16, 24 and 28 respectively. We notice that 25% of these papers use all three features while 30% of these papers use contextual and demographic information for extremist content detection. There are very few papers using only link information and statistical data for online radicalization detection.

Similar to Table 2, we classify articles into 3 subcategories i.e. 'content', 'user profile' and 'community' based upon the type of content being analyzed for providing solutions to counter and combat online extremism. Figure 12 shows a Venn diagram for number of articles classified in each category. Figure 12 reveals that out of 37 articles, in 35 papers, authors conduct experiments on contextual metadata while in 18 papers, authors extract user profile information. Similarly, in 16 articles, they mine linked user profiles and their communities.

As shown in Figure 12, content analysis is an important feature for identifying extremist content. Presence of multi-lingual text in contextual metadata makes it technically challenging to mine and analyze the data. Based upon our analysis, we classify articles into three categories based upon the language being addressed in experimental setup. Figure 13 shows that among 37 articles, 14 articles are capable to mine and identify extremist content posted in Arabic language. Among
these 14 articles, methods proposed in 7 articles are capable to analyze other non-English language texts as well. Figure 13 also reveals that all proposed methodologies in previous researches are capable to detect extremist content posted in English language.

As discussed in Section 5.3, many of the articles in online radicalization have been published in counter-terrorism and deradicalizing extremism. The focus of these researches is to counter and combat online radicalization targeting a specific country and originated from a specific region. Based upon the target region and country we classify these 37 articles into three categories: US Domestic, International and Others. Figure 14 shows the Venn diagram of number of articles classified in each category. Figure 14 reveals that there are 20 articles focusing on countering online radicalization happening worldwide. 5 articles focus on extremist content targeting issues of US domestic and originated from US domestic regions. In existing literature we find 9 such articles that focus on other extremist groups and regions worldwide. For example, North Africa, Latin America, Middle Eastern countries. We also find 3 articles targeting both US domestic, Middle Eastern and other regions for detecting online radicalization.

There are various genre of radicalization present on Internet. Based upon previous researches we divide genre into 5 categories: Anti-black or white supremacy communities are a group of people targeting non-whites. Anti-black is a form of racism spreading and promoting the ideology that white people are superior in certain characteristics, traits, and attributes to other non-white people. Jihad or Islamic extremism communities are the groups promoting Islamic extremism on social media. Act of Islamic extremism includes promoting "Sharia law" and abusing human rights and
terrorism (However, here we use terrorism as an independent category for classification). Anti-religion radicalization is posting hateful speech against a religion and promoting the beliefs and ideology of oneself. Hate and Extremism includes the others categories that are undefined and covers political radicalization. Figure 15 shows the number of articles classified in each genre. Due to the large number of articles in online radicalization domain, we use timeline on x-axis and number of articles on y-axis. Figure 15 reveals that over past decade there has been a constant research in each category of online radicalization. However, we observe that through out this timeline maximum number of publications are in the category of Jihad and terrorism. In early 20s there has been some work in the domain of white supremacy community detection on social media but after 2008, we find a sudden increase in the number of publication focusing on anti-black communities (13 publications). Figure 15 also reveals that the number of publications focusing on anti-religion based content are very few in comparison to other categories of online radicalization.

As discussed in previous Section (5.3), Precision, Recall, K-cross validation are standard information retrieval measures to evaluate the effectiveness of proposed solution approach. Based upon the literature articles of online radicalization detection we divide Evaluation dimension into 8 subcategories (illustrated in Table 4). F-score evaluates the harmonic mean of Precision and Recall while Accuracy computes the “correctness” of proposed method. Social Network Analysis (SNA) is an intuitive based technique to uncover patterns and relationships between entities in a network. User based analysis are the evaluations performed by external users. Figure 16 demonstrates the distribution of number of articles using various measures to examine the performance of proposed methods over past decade. As shown in Figure 12, 50% of the articles uncover the hidden communities of hate and extremist users. We find a similar pattern in Figure 16 which reveals that Social network analysis is the most commonly used technique to evaluate the performance of proposed method. Figure 16 also reveals that there is only one article using K-cross validation for evaluation while in 7 articles no evaluation method is defined. We also observe that despite having 50% of the articles that analyze only textual data to identify extremist content, only 30% of the articles use precision while only 20% of the articles use recall as a measure for evaluation.

6 Closing Remarks
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. We observe a surge in research interest over the last 3 years on the topic of solutions for identifying and forecasting civil unrest and mobilization by mining textual content in open-source social media. The number of research papers on social media analytics for online radicalization detection are much more (about 7 times) than on civil disobedience detection. Intelligence and Security Informatics (ISI) conference and Security Informatics (SI) journal are the two main venues for publishing papers on the topic of online radicalization detection and mining. Our analysis reveals that micro-blogging websites like Twitter and Tumblr are the two most common sources of social media data for civil unrest detection and forecasting applications. We believe that Twitter has been very instrumental in facilitating political mobilization in comparison to other social media platforms because of its inherent characteristics of sharing short text through direct messages and follower relationship. It is interesting to observe that despite the immense popularity and penetration of YouTube as an online video-sharing website, it has not been used in any of the existing research for protest planning or prediction. On the contrary, our research reveals that YouTube is the most widely used platform for online radicalization, hate and extremism promotion as indicated by the published research papers. In comparison to Twitter, Tumblr which is also a popular micro-blogging website has not been a major focus of research attention for online radicalization detection applications.

Our analysis reveals a variety of information retrieval and machine-learning based methods and techniques used by researchers to investigate solutions for civil unrest and online radicalization detection. Clustering, Logistic Regression and Dynamic Query Expansion are the commonly used techniques to predict upcoming events related to civil unrest or protest. We observe that Named Entity Recognition (NER) is a common component in the text processing pipeline for various proposed approaches and techniques. Graph modeling is also a technique adopted by several researchers for the problem of event forecasting. Our survey reveals that KNN (K Nearest Neighbor), Naive Bayes, Support Vector Machine, Rule Based Classier, Decision Tree, Clustering (Blog Spider), Exploratory Data Analysis (EDA), Topical Crawler/Link Analysis (Breadth First Search, Depth First Search, Best First Search) and Keyword Based Flagging (KBF) are the most widely used techniques for online radicalization detection on social media websites.

In this survey, we categorize existing studies on various dimensions such as the use of discriminatory features, type of metadata being analyzed and evaluation
techniques used by authors to examine the effectiveness of their approach. Our analysis reveals missing contextual metadata of a post and spatiotemporal information present in the content are most commonly used feature for predicting civil unrest related events. In existing studies, we find that maximum researchers evaluate the accuracy of their prediction approach by computing precision of their results. We also observe that 90% of the studies are able to mine English language text. Since the problem of civil unrest related event prediction can be targeted to a specific country or region, we find that the methods proposed in various studies are able to mine information from multi-lingual texts (Dutch, Spanish, French etc). Our analysis also reveals that 60% of the studies targets events specific to a country or region. Maximum number of studies are conducted on the events happened in Latin America and USA.

Characterization and meta analysis performed on the existing studies of online radicalization detection reveals that to identify the presence of extremist content contextual based metadata is most commonly used feature. However, demographic information and activity feeds of a user profile and links between two users are discriminatory features for locating hidden communities of extremist users. We observe that many of the existing techniques are capable to mine multi-lingual text such as Arabic and capture relevant information. Our analysis reveals that there are papers that targets only a country and region specific radicalization (Latin America, Middle Eastern, North Africa). We also observe that to examine the effectiveness of results in community detection and extremism content identification, Social Network Analysis and Precision measures are most commonly used evaluation methods.

References
1. Muthiah, S., Huang, B., Arredondo, J., Mares, D., Getoor, L., Katz, G., Ramakrishnan, N.: Planned Protest Modeling in News and Social Media
2. Ramakrishnan, N., Butler, P., Muthiah, S., Self, N., Khandpur, R., Saraf, P., Wang, W., Cadena, J., Vallikanti, A., Korkmaz, G., Kuhlman, C., Marathe, A., Zhao, L., Hua, T., Chen, F., Lu, C.T., Huang, B., Srivinas, A., Trinh, K., Getoor, L., Katz, G., Doyle, A., Ackermann, C., Zavorin, I., Ford, J., Summers, K., Fayed, Y., Arredondo, J., Gupta, D., Mares, D.: ‘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 ’14, pp. 1799–1808. ACM, New York, NY, USA (2014). doi:10.1145/2623330.2623373.
3. Filchenkov, A.A., Azarov, A.A., Abramov, M.V.: What is more predictable in social media: Election outcome or protest action? In: Proceedings of the 2014 Conference on Electronic Governance and Open Society: Challenges in Eurasia. EGOSE ’14, pp. 157–161. ACM, New York, NY, USA (2014). doi:10.1145/2729104.2729135.
4. Agarwal, S., Sureka, A.: A focused crawler for mining hate and extremism promoting videos on youtube. In: Proceedings of the 25th ACM Conference on Hypertext and Social Media, pp. 294–296 (2014). ACM
5. McNamee, L.G., Peterson, B.L., Peñ, J.: A call to educate, participate, invite and indict: Understanding the communication of online hate groups. Communication Monographs 77(2), 257–280 (2010).
6. Sureka, A., Agarwal, S.: Learning to classify hate and extremism promoting tweets. In: Intelligence and Security Informatics Conference (JISIC), 2014 IEEE Joint, pp. 320–320 (2014). IEEE
7. Chris Hale, W.: Extremism on the world wide web: a research review. Criminal Justice Studies 25(4), 343–356 (2012)
8. Fisher, A.: How jihadist networks maintain a persistent online presence. Perspectives on Terrorism 9(3) (2015)
9. Wang, M.C.G.A., Chen, Z.X.H., Mao, D.Z.W.: Intelligence and security informatics (2011)
10. Yannakogeoros, P.: Rethinking the threat of cyberterrorism. In: Chen, T.M., Jarvis, L., Macdonald, S. (eds.) Cyberterrorism, pp. 43–62. Springer, ??? (2014). doi:10.1007/978-1-4939-0962-9_3
11. Rowe, M., Stankovic, M., Dadzie, A.-S. (eds.): A Topical Crawler for Uncovering Hidden Communities of Extremist Micro-Bloggers on Tumblr (2019)
12. Zhao, L., Chen, F., Dai, J., Hua, T., Lu, C.T., Ramakrishnan, N.: Unsupervised spatial event detection in targeted domains with applications to civil unrest modeling. PLoS ONE 9(10), 110206 (2014). doi:10.1371/journal.pone.0110206
13. Naveed, N., Gorrion, T., Kunegis, J., Alhadi, A.C.: Bad news travel fast: A content-based analysis of interestingness on twitter. In: Proceedings of the 3rd International Web Science Conference, p. 8 (2011). ACM
14. Korkmaz, G.: Challenges of twitter-based predictions of civil unrest in latin america
15. Compton, R., Lee, C., Xu, J., Artimea-Moncada, L., Lu, T.-C., Silva, L., Macy, M.: Using publicly visible social media to build detailed forecasts of civil unrest. Security Informatics 3(1) (2014). doi:10.1186/s13388-014-0004-6
16. Xu, J., Lu, T.-C., Compton, R., Allen, D.: Civil unrest prediction: A twitter-based exploration. In: Kennedy, W., Agarwal, N., Yang, S. (eds.) Social Computing, Behavioral-Cultural Modeling and Prediction 2014. Lecture Notes in Computer Science, vol. 8393, pp. 403–411. Springer
17. Chen, F., Neill, D.B.: Non-parametric scan statistics for event detection and forecasting in heterogeneous social media graphs. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD ’14, pp. 1166–1175. ACM, New York, NY, USA (2014). doi:10.1145/2623330.2623619.
18. Hua, T., Lu, C.T., Ramakrishnan, N., Chen, F., Arredondo, J., Mares, D., Summers, K.: Analyzing civil unrest through social media. Computer 46(12), 80–84 (2013). doi:10.1109/MC.2013.442.
19. Colbaugh, R., Glass, K.: Early warning analysis for social diffusion events. Security Informatics 1(1), 1–26 (2012)
20. Leets, L.: Responses to internet hate sites: Is speech too free in cyberspace? Communication Law & Policy 6(2), 287–317 (2001)
21. Schafer, J.A.: Spinning the web of hate: Web-based hate propagation by extremist organizations. Journal of Criminal Justice and Popular Culture 9(2), 69–88 (2002)
22. Michael, G.: Confronting Right Wing Extremism and Terrorism in the USA. Routledge, ??? (2003)
23. Salem, A., Reid, E., Chen, H.: Content analysis of jihadi extremist groups’ videos. In: Intelligence and Security Informatics, pp. 615–620. Springer, ??? (2006)
24. Yang, C.C., Liu, N., Sageman, M.: Analyzing the terrorist social networks with visualization tools. In: Intelligence and Security Informatics, pp. 331–342. Springer, ??? (2006)
25. Zhou, Y., Qin, J., Lai, G., Reid, E., Chen, H.: Exploring the dark side of the web: Collection and analysis of us extremist online forums. In: Intelligence and Security Informatics, pp. 621–626. Springer, ??? (2006)
26. Korkmaz, G.: Challenges of twitter-based predictions of civil unrest in latin america
27. Compton, R., Lee, C., Xu, J., Artimea-Moncada, L., Lu, T.-C., Silva, L., Macy, M.: Using publicly visible social media to build detailed forecasts of civil unrest. Security Informatics 3(1) (2014). doi:10.1186/s13388-014-0004-6
26. Zhang, Y., Zeng, S., Fan, L., Dang, Y., Larson, C.A., Chen, H.: Dark web forums portal: Searching and analyzing jihadist forums. In: Proceedings of the 2009 IEEE International Conference on Intelligence and Security Informatics. ISI’09, pp. 71–76. IEEE Press, Piscataway, NJ, USA (2009). http://dl.acm.org/citation.cfm?id=1706428.1706441

27. Glass, K., Colbaugh, R.: Estimating the sentiment of social media content for security informatics applications. Security Informatics 1(1), 1–16 (2012)

28. Scanlon, J.R., Gerber, M.S.: Automatic detection of cyber-recruitment by violent extremists. Security Informatics 3(1), 1–10 (2014)

29. Bermingham, A., Conway, M., McInerney, L., O’Hare, N., Smeaton, A.F.: Combining social network analysis and sentiment analysis to explore the potential for online radicalisation. In: Social Network Analysis and Mining. 2009. ASOMAN’09. International Conference on Advances In, pp. 231–236 (2009). IEEE

30. Fu, T., Huang, C.-N., Chen, H.: Identification of extremist videos in online video sharing sites. In: Intelligence and Security Informatics, 2009. ISI’09, pp. 179–181 (2009). doi:10.1109/ISI.2009.5137295

31. Anand, T., Padmapriya, S., Kirubakaran, E.: Terror tracking using advanced web mining perspective. In: Intelligent Agent & Multi-Agent Systems. 2009. IAMA 2009. International Conference On, pp. 1–4 (2009). IEEE

32. Hovakimian, J., Wang, Y.: Locate the hate: Detecting tweets against blacks. In: AAAI (2013)

33. Wadhwa, P., Bhatia, M.: Tracking on-line radicalization using investigative data mining. In: Communications (NCC), 2013 National Conference On, pp. 1–5 (2013). IEEE

34. O’Callaghan, D., Greene, D., Conway, M., Carthy, J., Cunningham, d. Pá: Uncovering the wider structure of extreme right communities spanning popular online networks. In: Proceedings of the 5th Annual ACM Web Science Conference. pp. 276–285 (2013). ACM

35. O’Callaghan, D., Greene, D., Conway, M., Carthy, J., Cunningham, d. Pá: An analysis of interactions within and between extreme right communities in social media. In: Ubiquitous Social Media Analysis, pp. 88–107. Springer, ??(2013)

36. Berger, J., Strathern, B.: Who Matters Online: Measuring Influence, Evaluating Content and Countering Violent Extremism in Online Social Networks. (2013)

37. Botha, A.: Assessing the vulnerability of kenyan youths to radicalisation and extremism. Institute for Security Studies Papers (245), 28 (2013)

38. Blanquart, G., Cook, D.M.: Twitter influence and cumulative perceptions of extremist support: A case study of geet warriors (2013)

39. Chau, M., Xu, J.: A framework for locating and analyzing hate groups in blogs. PACIS 2006 Proceedings, 60 (2006)

40. Xu, J., Chau, M.: Mining communities of bloggers: A case study on cyber-hate. ICIS 2006 Proceedings, 11 (2006)

41. Ressler, S.: Social network analysis as an approach to combat terrorism: Past, present, and future research. Homeland Security Affairs 2(2), 1–10 (2006)

42. Davis, B.R.: Ending the cyber jihad: Combating terrorist exploitation of the internet with the rule of law and improved tools for cyber police. CommLaw Consentus 15, 119 (2006)

43. Conway, M., McInerney, L.: Jihad video and auto-radicalisation: Evidence from an exploratory youtube study. In: Ortiz-Arroyo, D., Larsen, H., Zeng, D., Hicks, D., Wagner, G. (eds.) Intelligence and Security Informatics. Lecture Notes in Computer Science, vol. 5376, pp. 108–118. Springer, ??(2008). doi:10.1007/978-3-540-89900-6_13 http://dx.doi.org/10.1007/978-3-540-89900-6_13

44. Sureka, A., Kumaraguru, P., Goyal, A., Chhabra, S.: Mining youtube to discover extremist videos, users and hidden communities. In: Cheng, P.-J., Kan, M.-Y., Lam, W., Nakov, P. (eds.) Information Retrieval Technology. Lecture Notes in Computer Science, vol. 6458, pp. 13–24. Springer, ??(2010). doi:10.1007/978-3-642-17187-1_2 http://dx.doi.org/10.1007/978-3-642-17187-1_2

45. Van Zoonen, L., Vis, F., Mihelj, S.: Performing citizenship on youtube: Activism, satire and online debate around the anti-islam video fitna. Critical Discourse Studies 7(4), 249–262 (2010)

46. Weimann, G.: New terrorism and new media. Wilson Center Common Labs (2014)

47. Chau, M., Xu, J.: Mining communities and their relationships in blogs: A study of online hate groups. International Journal of Human-Computer Studies 65(1), 57–70 (2007)

48. Fu, T., Abbasi, A., Chen, H.: A focused crawler for dark web forums. Journal of the American Society for Information Science and Technology 61(6), 1213–1231 (2010)

49. Danowski, J.A.: Changes in muslim nations’ centrality mined from open-source world jihad news: A comparison of networks in late 2010, early 2011, and post-bin laden. In: Intelligence and Security Informatics Conference (EISIC), 2011 European, pp. 314–321 (2011). doi:10.1109/EISIC.2011.69

50. Chen, H., Thoms, S., Fu, T.: Cyber extremism in web 2.0: An exploratory study of international jihadist groups. In: Intelligence and Security Informatics, 2008. ISI 2008. IEEE International Conference On, pp. 98–103 (2008). IEEE

51. Agarwal, S., Sureka, A.: Using knn and svm based one-class classifier for detecting online radicalization on twitter. In: Distributed Computing and Internet Technology, pp. 431–442. Springer, ??(2015)

52. Yang, M., Chen, H.: Partial clustering for radical opinion identification in hate group web forums. In: Intelligence and Security Informatics (ISI), 2012 IEEE International Conference On, pp. 96–101 (2012). doi:10.1109/ISI.2012.6284099

53. Ritter, A., Clark, S., Etzioni, O., et al.: Named entity recognition in tweets: an experimental study. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 1524–1534 (2011). Association for Computational Linguistics

54. Mahmood, S.: Online social networks: The overt and covert communication channels for terrorists and beyond. In: Homeland Security (HST), 2012 IEEE Conference on Technologies For, pp. 574–579 (2012). IEEE

55. Agarwal, S., Sureka, A.: Topic-specific youtube crawling to detect online radicalization. In: Databases in Networked Information Systems, pp. 133–151. Springer, ??(2015)

56. Sureka, A.: 140 characters of @ hate and protest (2012)

57. Wilner, A.S., Dubouloz, C.-J.: Transformative radicalization: Applying learning theory to islamist radicalization. Studies in Conflict & Terrorism 34(5), 418–438 (2011)

58. Zeng, D., Chen, H., Lusch, R., Li, S.-H.: Social media analytics and intelligence. Intelligent Systems, IEEE 25(6), 13–16 (2010)

59. Meddaugh, P.M., Kay, J.: Hate speech or “reasonable racism?” the other in stormfront. Journal of Mass Media Ethics 24(4), 251–268 (2009)

60. Brown, C.: Wwvw. hate. com: White supremacist discourse on the internet and the construction of whiteness ideology. The Howard Journal of Communications 20(2), 189–208 (2009)

61. Daniels, J.: Cloaked websites: propaganda, cyber-racism and epistemology in the digital era. New Media & Society 11(5), 659–683 (2009)

62. Qi, X., Christensen, K., Duval, R., Fuller, E., Spahiu, A., Wu, Q., Zhang, C.-Q.: A hierarchical algorithm for clustering extremist web pages. In: Advances in Social Networks Analysis and Mining (ASONAM), 2010 International Conference On, pp. 458–463 (2010). doi:10.1109/ASONAM.2010.81

63. van Zoonen, L., Vis, F., Mihelj, S.: Youtube interactions between agonism, antagonism and dialogue: Video responses to the anti-islam film fitna. new media & society 13(8), 1263–1300 (2011)

64. Rotman, D., Golbeck, J., Preece, J.: The community is where the rapport is – on sense and structure in the youtube community. In: Proceedings of the Fourth International Conference on Communities and Technologies. C&T#38’T 09, pp. 41–50. ACM, New York, NY, USA (2009). doi:10.1145/1556460.1556467

65. Correa, D., Sureka, A.: Solutions to detect and analyze online radicalization: a survey. arXiv preprint arXiv:1301.4916 (2013)

66. Stern, J.: Mind over martyr: How to deradicalize islamist extremists. Foreign Affairs, 95–108 (2010)

67. Qin, J., Zhou, Y., Reid, E., Lai, G., Chen, H.: Analyzing terror campaigns on the internet: Technical sophistication, content richness, and web interactivity. International Journal of Human-Computer Studies 65(1), 71–84 (2007)

68. Zhou, Y., Reid, E., Qin, J., Chen, H., Lai, G.: Us domestic extremist
groups on the web: link and content analysis. Intelligent Systems, IEEE 20(5) (2005)

69. Smith, A.G., Suedfeld, P., Conway III, L.G., Winter, D.G.: The language of violence: Distinguishing terrorist from nonterrorist groups by thematic content analysis. Dynamics of Asymmetric Conflict 1(2), 142–163 (2008)

70. Reid, E., Chen, H.: Internet-savvy us and middle eastern extremist groups. Mobilization: An International Quarterly 12(2), 177–192 (2007)

71. Erez, E., Weimann, G., Weisburd, A.: Jihad, crime and the internet: Content analysis of jihadist forum discussions. Final Report submitted to the National Institute of Justice in fulfillment of requirements for Award (2006-IJ) (2011)

72. Reid, E., Qin, J., Zhou, Y., Lai, G., Sageman, M., Weimann, G., Chen, H.: Collecting and analyzing the presence of terrorists on the web: A case study of jihad websites. In: Intelligence and Security Informatics, pp. 402–411. Springer, ??? (2005)

73. Park, A.J., Tsang, H.H., Sun, M., Gläser, U.: An agent-based model and computational framework for counter-terrorism and public safety based on swarm intelligence. Security Informatics 1(1), 1–9 (2012)

74. Hawdon, J.: Applying differential association theory to online hate groups: a theoretical statement (2012)

75. Davis, P.K.: Influencing violent extremist organizations and their supporters without adverse side effects (2012)

76. Iannaccone, L.R., Berman, E.: Religious extremism: The good, the bad, and the deadly. Public Choice 128(1-2), 109–129 (2006)

77. Akoumianakis, D., Kafousis, I., Karadimitriou, N., Tsiknakis, M.: Retaining and exploring online remains on youtube. In: Emerging Intelligent Data and Web Technologies (EIDWT), 2012 Third International Conference On, pp. 89–96 (2012). IEEE

78. Hwang, J.C.: Terrorism in perspective: An assessment of ‘jihad project’ trends in india. Asia Pacific Issues (104) (2012)

79. Torres Soriano, M.R.: The road to media jihad: the propaganda actions of al Qaeda in the Islamic maghreb. Terrorism and Political Violence 23(1), 72–88 (2010)

80. Lugna, L.: Institutional framework of the european union counter-terrorism policy setting. Baltic Security & Defence Review 8(1), 101–128 (2006)

81. Bhattacharjee, J.: Understanding 12 extremist groups of bangladesh. Observer India (2009)

82. Barnett, B.A.: Hate group community-building online: A case study in the visual content of internet hate sites. Proceedings of the New York State Communication Association (2007)

83. Braha, D.: Global civil unrest: contagion, self-organization, and prediction. PloS one 7(10), 48596 (2012)

84. Compton, R., Lee, C., Lu, T.-C., de Silva, L., Macy, M.: Detecting future social unrest in unprocessed twitter data: #x201c;emerging phenomena and big data #x201d;: In: Intelligence and Security Informatics (ISI), 2013 IEEE International Conference On, pp. 56–60 (2013). doi:10.1109/ISI.2013.6578786

85. Doyle, A., Katz, G., Summers, K., Ackermann, C., Zavorin, I., Lim, Z., Muthiah, S., Zhao, L., Lu, C.-T., Butler, P., et al.: The embers architecture for streaming predictive analytics. In: Big Data (Big Data), 2014 IEEE International Conference On, pp. 11–13 (2014). IEEE

86. Crowley, D.N., Passant, A., Breslin, J.G.: Short paper: annotating microblog posts with sensor data for emergency reporting applications. In: Proceedings of the 4th International Workshop on Semantic Sensor Networks 2011 (SSN11), a Workshop of the 10th International Semantic Web Conference-ISWC 2011: 23-27 October 2011; Bonn, Germany, pp. 84–89 (2011)