Geosocial Media as a Proxy for Security: A Review

ZHIGANG HAN [1,2,3,4,5]**, SONGNIAN LI**[2]**, CAIHUI CUI**[3,5]**, DAOJUN HAN**[6]**, AND HONGQUAN SONG**[1,3,5]

[1]Key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions, Ministry of Education, Kaifeng 475004, China
[2]Department of Civil Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada
[3]College of Environment and Planning, Henan University, Kaifeng 475004, China
[4]Henan Industrial Technology Academy of Spatio-Temporal Bigdata, Henan University, Zhengzhou 450008, China
[5]Urban Big Data Institute, Henan University, Kaifeng 475004, China
[6]School of Computer and Information Engineering, Henan University, Kaifeng 475004, China

Corresponding author: Songnian Li (snli@ryerson.ca)

ABSTRACT Security issues such as natural disasters and terrorist attacks have attracted increasing global concern and attention. How to effectively detect security events has become worrisome to countries worldwide. Advances in mobile Internet technology have led to hundreds of millions of users using social media daily to post microblogs, text messages and multimedia information, generating enormous amounts of social media data that reflect people's social behaviors. Studies have proven that the timely intelligence can be extracted from these data. In particular, geosocial media when combined with location information can be used as a proxy for security event detection and security situational awareness. This paper provides a synopsis of the geosocial media data and the related processing/analysis methods used for detecting security events, and summarizes the general framework of security-related analyses based on geosocial media. Four major categories of analysis methods and application cases, including natural language processing, social network analysis, location inference and geospatial analysis, and image or video understanding, are discussed in detail. The paper concludes with possible future directions and areas of research that could be addressed and investigated. We hope to provide a clarion call to the scientists, practitioners, and other stakeholders to enhance the capabilities and accuracy of security events detection and security situational awareness and assessment using geosocial media.

INDEX TERMS Geosocial media, security, geospatial, event detection, situation awareness.

I. INTRODUCTION

In recent years, with the worsening of global warming, religious problems and imbalances in social development, security incidents have increased substantially, causing anxiousness among the majority of the world's population and causing them to worry about their safety and wellbeing. In Japan and some coastal areas of the United States and Australia, natural disasters caused by climate change, such as floods and tsunamis, have had substantial impacts on people's productivity and lives. In some areas in the Middle East

The associate editor coordinating the review of this manuscript and approving it for publication was Kaitai Liang.

and Africa, people live in war zones and face all types of threats, including death or exile, while in the relatively more affluent areas of the world, people are fearful of terrorist attacks, murders, riots and other criminal incidents. A sense of insecurity is pervasive worldwide - we are living in an era of insecurity [1]. These security issues have a grave impact on the economic and social development of countries around the world every year. For example, approximately 9,000 people were killed by disasters in the forms of floods, forest fires and earthquakes, and 96 million people were affected by disasters, which caused a combined loss of 270 billion euros in 2017, the second highest annual loss in history [2]. Therefore, effective and timely methods to detect and respond to security

events to minimize the losses they cause have become a focus of the United Nations, governments, enterprises, nongovernmental organizations (NGOs) and individuals.

With the rapid development of mobile Internet technology, social media has emerged and rapidly become a part of people's daily lives. Because access to the Internet can easily be achieved through a smartphone, people can post and share information on social media about their experiences, opinions or views, possibly tagged with their location, at any time. Social media is a type of Internet application based on web, and mobile terminals support creating, accessing, and exchanging of various pieces of information in the form of user-generated content (UGC) [3], [4]. Social media involves a range of aspects, such as communication, community, connectivity, collaboration, and UGC generation and sharing. Social media applications are free and they make it easy for people to socialize, allowing users to upload what they see and think at any time in various forms (e.g., text, image and video) [5]. Most social media technologies originated before 2005 when Tim Berners-Lee established the Web 2.0 concept. Subsequently, Wikipedia, blogs, social networking sites, Weibo and other content-sharing sites have become extremely popular [6].

Social media has the advantages of being low cost and near real-time with the user participation characteristics of Web 2.0. It has also become a powerful tool for commercial marketing and business promotion. Currently, hundreds of millions of interactive content entries are being published on major social media platforms such as Facebook, Twitter, and online forums by a massive number of users. These interactive content entries have also become a critical data sources for global scientific research, especially in the field of computable social science field [7]. However, due to the rapid spread and diffusion of information on social media platforms, unpredictable consequences can occur in the absence of a timely response. The Arab Spring is a typical example. News of the self-immolation incident of a Tunisian peddler that occurred in late 2010 spread rapidly on social media platforms, which triggered a wave of protests in Tunisia against the Tunisian government that quickly expanded to neighboring countries such as Libya. Social media was one of the unique factors driving this event via video and photo sharing and by people messaging and forwarding text messages related to the protests.

Social media data provide rich information that reflects people's social behavior. In the security field, various groups of terrorists and gangs have increasingly recognized the value of social media and have actively used it to plan and organize activities, recruit members, spread terrorist ideas and publish various terrorist messages to expand their influences. Some extremist organizations have established their official accounts on platforms such as Twitter; others have uploaded videos on YouTube to recruit members and conduct terrorist attacks. A typical case is the recruitment of new members such as Arid Uka for terrorist organizations and getting them to perform terrorist actions. Such actions form new threats

faced by governments today [8]. Social media has magnified the influence of terrorist groups and enhanced their ability to organize terrorist activities. The connectivity characteristics of social media have been used to organize and collaborate activities to oppose governments or military and police organizations. Terrorist recruitment, propaganda and information dissemination are also utilized the broadcast characteristics of social media [9].

Security incidents may occur at different times and in different locations, involving different people and objects. These elements contain the spatiotemporal information of security events, related users, their relationships and other objective information: these are the essential information for perceiving security events and assessing security situations. Traditionally, access to such information has relied on professional intelligence channels. However, social media applications, such as instant messaging platforms running on the Internet, have large numbers of users (including various terrorist organizations and gangs) who generate a steady streams consisting of large amounts of unstructured data, including texts, videos, images and audio, as well as a large amounts of user interaction data (e.g., social network data such as views and forwarding). These massive amounts of data include various pieces of information related to security incidents; thus, fully mining and utilizing these data to perceive and assess security problems and respond to emergencies is of great practical significance. However, until recently, the field of security lacked systematic review and examination of such social media applications. In this study, to provide support for deepening the application of social media in the security field, we focus on real-world security issues and attempt to systematically examine the applications of geosocial media data, particularly those with location tags, in the context of security event detection and the assessment and perception of security situations.

The goals of this review are to identify datasets, methods and use cases utilizing geosocial media for security-related analysis tasks. We queried the eligible literature sources from the literature search portals such as the Clarivate Web of Sciences, Google Scholar and the digital libraries of organizations and publishers such as IEEE, ACM, Elsevier, Wiley, Taylor & Francis, Sage, MDPI, and Springer. The main works considered included books, journals, and the proceedings of the workshops and conferences published between 2008 and 2018 in English. From this search, 256 papers were initially identified. Then, screening was performed by reading the full-text, and examining methods and use cases. We primarily focused on papers that clearly described the methods used concerning the exploration, extraction, processing, validation, and aggregation of ***geosocial media data*** for security purposes. After removing the duplicate and weakly relevant content, 85 papers were selected that showed higher relevance to the security-related analysis tasks using geosocial media data. Furthermore, the prior references were searched to expand the background and current progress of the field. Finally, we used 154 literature records in this review process.

TABLE 1. Taxonomy of security-related analysis tasks.

Security tasks		Related papers and its topics
Security Events Detection and Sensing	Natural disasters	Earthquake [10-15]; Hurricane [13,15-18]; Flood [14,19]; Climate [20,21]; Wild fire [14]
	Man-made disasters	Oil spills [11,22]; Urban fires [23-25]; Nuclear disasters [26]
	Violence	Urban hijacking [23,27]; Terrorism [28-38]; Gang violence [39]; Bombings [40-43]; Violent Behavior [44-48]
	Military-related	Civil war [49]; Syrian war [50]; ISIS attacks [51,52]; Gaza conflict [53,54]; Israeli–Hamas conflict [55]
	Sociopolitical	Riots [15,56]; Protests [57,58]; Political [59,60]
	Others	Violent ideologies [61-67]; Threatening messages [68]; Hate speech [69-71]; Epidemics [72]; Radical [73]
Security Situational Awareness and Assessment		Geographic profiling of terrorist attacks [74]; Spatial and temporal patterns of terrorist attacks [75,76]; Geography of suicide terrorism in Israel [77]; Wealth dispersion and conflict location [78]; Risk assessment of terrorist attacks [79,80]; Spatializing the social networks of gang violence [81,82]

The remainder of this paper is organized as follows. Section 2 presents the security-related analysis tasks, Section 3 summarizes the framework for security-related analysis tasks and the location extraction or inference methods. Section 4 describes the platforms and data sources for geosocial media and data collection techniques. Section 5 provides the details of the key methods and technology for mining security information from geosocial media. Some observations are made and discussed in Section 6. As a complex task, the collaborative, real-time, and deep learning for security detection and awareness are also discussed. Section 7 concludes the paper with a brief summary.

II. SECURITY-RELATED ANALYSIS TASKS ON GEOSOCIAL MEDIA

Real-world security issues involve spatial locations. Consequently, we categorize the security-related analysis tasks into two groups with related papers and topics, as shown in Table 1. First, security events are detected and sensed at a place level, in which a specific place where an event occurs is treated as the space of the incident, emphasizing various security incidents that occur at a given scale within a short period of time, e.g., terrorist activities in city squares or streets and riots and protests in certain parts of a city. Second, security situations are identified and assessed at a regional scale, this discussion focuses on the perception and assessment of the overall security situation and the patterns of a particular city or region over longer temporal scale and a larger spatial scale, involving the overall security situation and pattern within a specified range.

A. SECURITY EVENT DETECTION AND SENSING

In general, an event refers to an incident with causal antecedents and consequences that is occurring or is about to occur at a particular moment and place [83]. Many types of security incidents exist in the security field, which may lead to varying degrees of devastating consequences depending on the scale of the incident place, the population affected and the level of panic caused. Atkinson *et al.* [84] grouped security events into six categories: natural disasters, man-made disasters, violent events, military events, social events, and political events. The classification defines the primary research subject of security events detection, i.e., using multisource data and multiple techniques to perceive the onset and occurrence of the aforementioned security events in a timely manner, identify high-risk groups and areas, track their development processes, predict their developing trends, apply appropriate intervention measures and respond as early as possible to minimize adverse impacts. From the perspective of time-effectiveness, security event detection includes NED (new event detection) and RED (retrospective news event detection) [85], [86]. NED primarily identifies security events from social media in real time, e.g., rapid tweet-based detection of earthquake events [10] or the detection of sudden crime and disease-related events [87], while RED attempts to identify previously unknown events from historically accumulated documents, e.g., the ETree system that tracks the evolution of events [88].

B. SECURITY SITUATION AWARENESS AND ASSESSMENT

Perceiving and assessing security situations on a regional scale, analyzing the spatial-temporal patterns of security events, and identifying their influencing factors are conducive to ensuring timely warnings and correct responses to security problems. The regional security situation is a quantitative assessment of the risk of security events at a particular scale. The security events can be single or composite types, and the assessment scale can be a country, region or specific city. For example, the field of terrorism research has undergone a geographical shift [89] and geography should provide more support in the security field [90]. Some scholars have used geographical methods to study the relationship between domestic and global terrorism [91], performed geographical profiling by combining locations of various terrorist attack incidents, identified their spatial patterns [74], assessed terrorism risks in a region, and analyzed the spatial distribution patterns of a region's vulnerability [79], [80]. The Institute for Economics and Peace (IEP) has designed a map of global terrorism indicators (GTI) to describe the main spatial distribution patterns and trends of global terrorism [92] (Figure 1). Similar cases include an analysis of the spatiotemporal agglomeration pattern of terrorist attacks [75], an assessment of the loss pattern of the Syrian civil war [93], and an investigation of the geographical factors of Palestinian suicide terrorism [77]. It is found that within a country,

FIGURE 1. Global Terrorism Index Map (Data source: Global Terrorism Index Annual Report in 2018).

FIGURE 2. A typical framework for security-related analysis tasks using geosocial media data.

regions with a large gap between the national average economic level and the lower regional net income level are more prone to violent internal conflicts [78].

III. A TYPICAL FRAMEWORK AND LOCATION EXTRACTION FOR SECURITY-RELATED ANALYSIS TASKS

A. A TYPICAL FRAMEWORK FOR SECURITY-RELATED ANALYSIS TASKS USING GEOSOCIAL MEDIA

We summarize the general process of security event detection and security situational awareness and assessment using geosocial media (Figure 2). First, based on the identification of social media platforms (data collection sources) and the keyword list screened from the corpus of the security field, the data related to specific keywords and the metadata are acquired from various social media platforms using various data collection methods. Next, data cleaning and filtering are conducted to generate structured datasets that are used as the basis for text or image feature extraction and social network construction that are performed. Location extraction and inference are then performed on the four types of data of text, social network, metadata, and image. Next, the text, graph, and image features are extracted to detect and analyze the security events or situation. These operations involve four types of technologies: (1) text analysis, including named entity recognition (NER) and text similarity calculation, topic modeling and sentiment analysis; (2) social network analysis, including user impact analysis, community discovery and complexity network computing; (3) geospatial analysis using multisource data; and (4) image comprehension, in which object recognition, video comprehension, and visual analysis are conducted using the collected social media data, such as images and short videos. On this basis, the 6W information of security events (i.e., who, which, what, when, where, and why) are identified, and various features of the security-related analysis tasks such as process and evolution, location and distribution, user and community, and subject and behavior are further detected, summarized and visualized to identify potential security issues and provide decision support to policymakers.
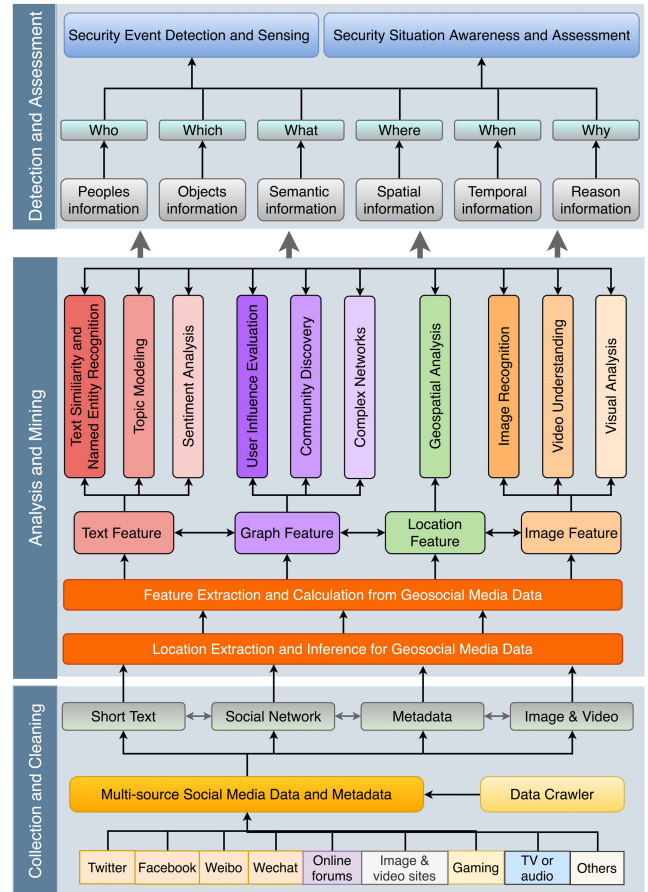
B. LOCATION EXTRACTION AND INFERENCE FOR GEOSOCIAL MEDIA

All security events occur in specific locations. It has been shown that location extraction is the foundation for security-related analysis tasks (Figure 2). Some social media content, albeit a small percentage, contain geotags from which geographic locations of users or messages can be extracted. For example, geotagged tweets occupy approximately 1 to 5% of the total number of tweets [94], and some have low accuracy. The majority of social media do not have geotags, thus, the location inference must be performed. In the early days of social media, location inference was conducted through Internet Protocol (IP) address-based approaches [95]. This approach is less accurate and relies on IP location libraries. More recently, location information is inferred based on text content and social network structure. Social media texts related to location information include various types, such as the user self-reported place names or locations, the user's description profile, the user name, the tweets themselves, or a combination of such data [96]. By performing NER on these text content entries, their location information can be inferred through algorithms such as Naive Bayes, AdaBoost, and logistic regression. SpinRec is a multistage collaborative filtering model to make location

TABLE 2. Taxonomy of data sources and platforms for security.

Data sets	Types	Data contents [a]					Platforms or datasets and paper numbers
		T	G	L	I	V	
Social Media Data Sets	social network sites	☑	☑	☑	☑	☑	Twitter(49);Facebook(2); Weibo(1);WeChat; Telgram; Tumblr
	image sharing sites		☑	☑	☑		Flickr(3); Instagram(1); Pinterest
	video sharing sites		☑	☑		☑	YouTube(2); Youku; Tiktok
	Forums or blogs	☑	☑	☑	☑		Reddit or the dark web(6)
	Game sites		☑			☑	Twitch; Steam
	others	☑		☑	☑		Ushahidi or RapidSMS, etc.(4)
Open Data Sets	Security database	☑		☑			GDELT, MOD, AUTO, etc. (17)
	Text corpus	☑					iWeb; WordNet; Google Web 1T
	Image/video labeled datasets				☑	☑	ImageNet; YouTube-8 M
	Geospatial datasets		☑	☑	☑		Geonames;OpenStreetMap; GDAM; GPW

[a]T is text, G is graph, L is location, I is image, and V is video. It means that the data sets have the corresponding data mainly if the table cell of data contents is checked.

inference based on a matrix decomposition method through the linkages between keywords used by people at different locations [21]. By detecting cooccurrence terms in tweet texts based on spatial point patterns and statistical methods, user location can be inferred without relying on any other external place-name database [97]. A knowledge-based joint neural network model for user location inference was also developed by combining a large-scale knowledge base with a text-based approach [98]. The follower and reposting networks of social networks can be used for location inference. A large-scale user location estimation system, named SPOT, can estimate social media users' locations based on features such as friends, social closeness and the local social coefficient [99]. It was found that the follower network approach could achieve higher accuracy and recall for social media location information inference [96].

IV. GEOSOCIAL MEDIA DATASETS AND DATA SOURCES

According to the application domain and goals, security-related analysis tasks require structured and unstructured data from many different sources. Text, image or video, and social network data are the fundamental sources. In addition, various types of auxiliary analysis data such as security databases, text corpora, annotations to images or videos, geospatial information, etc. are also indispensable for analysis (Table 2). These data may come from a variety of different platforms and require multiple data collection and cleaning methods to obtain complete, normalized datasets.

A. GEOSOCIAL MEDIA DATASETS
Social media provides timely first-hand information for the detection and perception of security events, and the latest news and observations are first broadcasted through social

media. For example, the news of Osama bin Laden's death was first reported by a Pakistani user on Twitter [83]. The data from these social media platforms are primarily of two types: UGC and metadata. The UGC can be of various forms such as short message text, photo/picture or audio/video. Due to the text length limitation, the majority of these short texts use a large number of abbreviations and emojis. A Crowded platform was developed to achieve the real-time perception of security events by collecting images of specific locations from online platforms [100]. Metadata consist of the data generated and released with the UGC, which, in addition to background context data such as release time and location, include the message reading data such as the number of reads, forwards and likes, as well as the social network and interaction data such as the views and forwards among users. UGC and metadata exist largely as semi-structured or unstructured data available to users through web searches and data APIs.

B. AUXILIARY ANALYSIS DATASETS
Security analyses also require a large number of auxiliary datasets. In general, four types of auxiliary analysis data exist. Security databases are mostly comprehensive datasets containing various security events. The most common is the Global Terrorism Database (GTD), which contains the details of more than 180,000 global terrorist attacks between 1970 and 2017. The Joint Research Center of the European Union has developed a multilanguage event extraction system and has built two types of security event datasets (MOD and AUTO) based on online news [84]. Many public corpora are available in the field of natural language comprehension, e.g., iWeb, WordNet, and the Embedding Corpus for Chinese Words and Phrases. When inferring the location of a corresponding event, geographic entity recognition must be performed from a gazetteer. GeoNames, a geographical name database, contains more than 25 million place names from countries around the world [101]. The OpenStreetMap dataset contains multiple feature layers such as roads, buildings and water bodies [102]. Other relevant geographic information data includes the Database of Global Administrative Areas (GDAM) and the Gridded Population of the World (GPW). Labeled image/video data and their pretrained models are also important auxiliary data; parameter adjustment and training can be conducted based on existing datasets and models to generate analysis results. For example, Google's newly released pretrained natural language processing (NLP) model BERT allows users to conduct NLP starting from a pretrained model to improve model accuracy [103].

C. HARNESSING GEOSOCIAL MEDIA DATA: CAPTURING, CLEANING AND PREPROCESSING
Social media data can be collected in various ways such as APIs, feeds and crawlers. An API is a data access interface provided by social media platforms. For example, Twitter provides two types of access APIs: search and streaming. In addition, commercial data providers such as Gnip and DataSift offer data purchasing services to meet analysis and
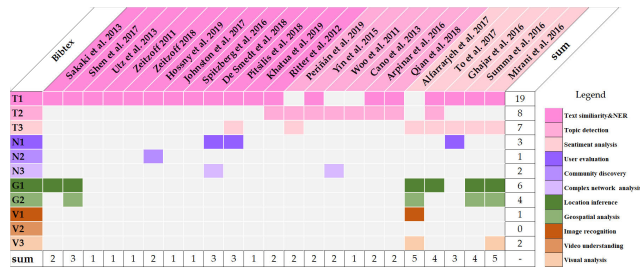
FIGURE 3. 23 selected papers of NLP for security using geosocial media. T1-T3: text similarity and NER, topic detection and sentiment analysis; N1-N3: user evaluation, community discovery, and complex network analysis; G1-G2: location inference, and geospatial analysis; V1-V3: image recognition, video understanding, and visual analysis.

application needs at different levels. RSS feeds is widely used in online news channels, blogs, and wikis. Indexing and collecting online news and blog content by reading RSS feeds is another important approach to social media data acquisition. Moreover, the web data can be collected by building a web crawler. One available open source crawler framework is Nutch; another widely used crawler is the distributed crawler framework Scrapy application built in Python.

To avoid the issues of "garbage in, garbage out" during analysis, preprocessing operations and data quality validation should be performed after the initial data collection. To achieve high-quality text mining, various data cleaning and preprocessing methods are required, such as spellchecking, deleting duplicate or abnormal characters, etc. These cleaning tasks can be done through text processing techniques using regular expressions or through data cleaning tools such as Google OpenRefine or RapidMiner. As the variety of geosocial media data, the data quality verification can strengthen the validity of the data, avoiding errors, over-presence or duplicated information. A more synthetic review of the state-of-the-art of geosocial media data quality was carried out in the literature [104]. After data cleaning, with the help of an NLP library such as NLTK or spaCy, a series of preprocessing procedures such as word segmentation, stop words removal, word stem completion, and part-of-speech tagging are applied to provide the foundation for feature extraction and analysis.

V. KEY APPROACHES AND TECHNOLOGIES USED IN THE ANALYSIS PROCESS

The critical technologies of security event detection and situational awareness in this section are described from four aspects: natural language processing, social network analysis, geospatial analysis, and image or video understanding and visual analysis. The methods are subdivided and the papers are categorized by the main method of study. If a paper involves multiple related technologies, it is given multiple color markers. To compare the different techniques, location inferences are also added to the figures (see Figure 3–6 in the following subsections).

A. NLP APPROACHES FOR TEXT CONTENT

Text entries are the basic source of security-related analytical tasks using geosocial media. Extracting security-related keywords from texts, analyzing their topics and evaluating their emotional tendencies are key aspects of social media-based security event analysis (Figure 3). Text similarity calculation and NER can be performed based on text keywords to find security-related information and events [105]. The similarity metrics such as Damerau-Levenshtein, Jaccard, and N-gram which are based on strings can be used to text similarity calculations [106]. By counting the numbers of keywords in tweets and using the Jaccard similarity indicator, riots can be detected in a city [56]. A tweet classifier based on features such as keywords and their numbers was used to develop a seismic event detection system [10]. Machine learning algorithms such as the support vector machine (SVM) can be used to automatically detect hate speech in Twitter [69]. Similar works include the recognition of Sunni extremist propaganda [67] and classification of tweets during an outbreak of the Ebola virus [72]. NER has been used to detect conceptual entities in a text including the names of people, places and organizations, which are the core elements of security event detection. In the early stages, NER used the linguistic grammars method to manually construct rule templates. Based on a combination of regular expression techniques, NER uses the matching of templates and strings as its main criterion. The conditional random field (CRF) is the main model currently used in NER [107]. Location entities can be extracted from Twitter by training a CRF model [108]. With the continuous development of neural networks, deep learning models such as the convolutional neural network (CNN) and the recurrent neural network (RNN)-CRF have been applied to NER [109], and attention mechanisms have been introduced into neural networks, generating better results [110].

Topic modeling can be used to extract topic information from unstructured text to perform topic clustering or text categorization. Currently, topic extraction mainly adopts probability-based models such as the Latent Dirichlet Allocation (LDA) [111]. Given the limited lengths of text messages combined with the user information and contextual features in social media such as hashtag, topic models for labeled LDA [112] and Twitter-LDA [113] have been defined for short texts such as tweets; improving the results for short text topic recognition. The environment-relevant topics were detected using CASPER system to analyze Spanish-language tweets [20]. A new method of hate speech classification named HCVAE was proposed to distinguish 40 hate subcategories of 13 different hate topic categories, which has raised tweet topic classification to a more detailed level [70]. Seven main topics were extracted from the Ummah dataset of the Dark Web Forum Portal for extreme ideologies diffusion on online forums [28]. A violence detection model (VDM) can identify texts containing violent content even in the absence of any tagged corpus and extract violence-related topics from social media data [44].
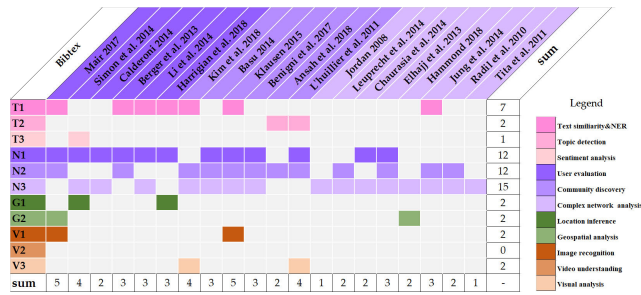
FIGURE 4. 20 selected papers of social network analysis for security using geosocial media.

Using sentiment analysis technology in NLP, it is possible to make a "positive," "neutral" or "negative" judgment of the emotional tendency of a text and track user's views on different topics [114]. Twitter has become the primary source of data for sentiment analysis for investigating how public emotions are affected by social, political, cultural and economic events [94]. Sentiment analysis is essentially a classification problem that includes supervised and unsupervised methods. In the supervised approaches, early studies categorized texts into various emotional categories using the annotated datasets by relying on the emotional tendencies of texts. The unsupervised methods use a sentiment lexicon to parse syntactic patterns [115]. A sentiment analysis on the Islamic State of Iraq and Syria (ISIS)-related tweets [51] or the tweets posted by ISIS Fangirls [31] were conducted. Another example is the sentiment evaluation of the tweets of the Boston Marathon bombing event [41]. Image data such as Flickr images can also be used for sentiment evaluation [13]. The studies such as Coooolll system [116] and unsupervised neural language model [117] were used to sentiment analysis based on deep learning techniques, which is the current focus of research.

In short, for the text content, the similarity and NER techniques are widely used to extract the security-related information, while topic modeling and sentiment analysis are used to reveal the topics and opinions of the texts on geosocial media. Machine learning methods such as SVM, CRF, and LDA are mainly used for NLP, and NER is used in all three analysis tasks. From the perspective of the method integration, the NLP method has less integration with others except for geospatial analysis. As the artificial neural network advances, deep learning methods have been gradually applied to NLP.

B. SOCIAL NETWORK ANALYSIS

A social network is often described by a graph structure consisting of major analytical dimensions such as users, communities and network structures (Figure 4). User's impact on a social network can be evaluated, and various indicators were systematically examined and classified for Twitter user's impact [118]. It is found that the betweenness centrality is the most important predictor of the key members of the Mafia organization [39]. The most influential social media accounts and those most prone to be influenced for white supremacists can be found by a scoring system [61]. An analysis of the characteristics of different users in social media groups helps to address security events reasonably [32]. By analyzing how the terrorist organization al-Shabaab used Twitter, it was found that al-Shabaab was mainly interested in describing the attack scene to attract its target audience [30]. In addition to influence, the analysis and identification of user profiles enables the acquisition of background information such as the user's education and place of residence. Text features and network information can identify user attribute parameters [119]. Location tags for Reddit users can be generated despite the absence of explicit geotagging data [120]. By mining the common characteristics of the same user from different networks, users can be matched in heterogeneous social networks. Multidimensional user features were extracted from multiple heterogeneous networks and used to identify the correspondence among different accounts belonging to the same user [121]. A deep neural network-based user identification algorithm named DeepLink was proposed to perform representative learning using social network sampling to achieve user identification [122].

Social network users have a stable group (community) structure of individuals and social relations [123], [124]. The two types of community structures are centered on either individuals or topics. The detection methods for a community structure mainly utilize the graph clustering and similarity calculation methods. After calculating node similarity, the network is divided according to that a higher the node similarity exists within the subnetwork, while a lower the number of connections exist between those nodes in different subnetworks. Another community detection idea is to reconstruct a network group based on the retention or deletion of the edges. Currently, the major algorithms include Girvan-Newman, fast greedy, fast unfolding, label propagation, and Infomap [125], which can use modularity indicators to evaluate community test results [126]. The community structure of social networks on Facebook during the 2016 Louisiana flood disaster was determined using the Girvan-Newman algorithm [16]. More than 200 terrorist attacks in India were analyzed to identify the responsible groups and their structures in the network [34]. Online extreme communities that support ISIS can be detected using iterative vertex clustering and classification (IVCC) approach [52]. Similar investigations have included identification of virtual communities and key members in the dark web [62], and a community analysis of Twitter to identify Western foreign fighters in Syrian and Iraq attacks [50]. Events can be detected based on changes in the community's size. In the SensorTree system, sudden protest events were detected based on the dramatic changes in the sizes of these communities [57].

Analyses of social networks topology helps in understanding the user roles and the community characteristics and their evolution [127]. Currently, complex networks have defined a series of indicators that measure the network centrality structure. A number of indicators such as network density were
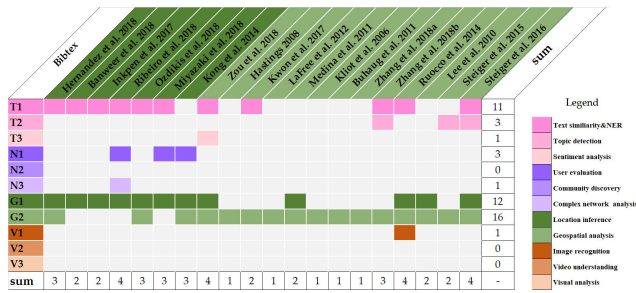
FIGURE 5. 20 selected papers of location inferences and geospatial analysis for security using geosocial media. Note: the location inference methods are summarized in Section 3.2.

used to analyze the risk communication network of Twitter users during the North Korean nuclear tests in 2013 [59]. The network centrality measurements of terrorists were used to analyze their social network characteristics [43]. A network centrality matrix was built to infer the functional characteristics of terrorist networks [33]. The centrality indicators of Lashkar-e-Taiba (LeT) terrorist social networks was calculated in the Mumbai attack incident [35]. Users in social networks are active in their corresponding spaces or places; thus, the embedded spatial information in social networks can help identify the spatial distribution characteristics of the networks. Combined the geospatial locations with those in social networks, a spatialization of social network analysis was achieved [81]. It was found that their correlations were stronger than those obtained by the neighborhood-based spatial autocorrelation measurements through analyzing the spatial distributions of gangs' competitive networks and criminal activities [82].

In addition to the social networks, the infrastructure network structures are worthy of attention. It was found that the attack locations were not the result of random selection through analysis of the locations of the subway station attacks in London and the structure of the subway network [29]. The military conflicts are more likely to occur in locations with a high degree of betweenness centrality in the road network, i.e., a gateway location that controls access to other areas [49].

To summarize, social networks are the essential feature for geosocial media; their users, communities, and network structures can be analyzed to reveal related users and its interactions on geosocial media. These complex network methods in are heavily used in user influence evaluation, community detection, and in investigating social network structures. The combination of geospatial and network spaces has drawn increasing attention in the security field. Because a given user may participate in different social networks, a user matching and identification method requires the feature mining from multiple social media platforms.

C. GEOSPATIAL ANALYSIS

The location is the basic information for event detection and the analysis of event spatial patterns (Figure 5). Based on the location information, the spatial agglomeration

and distribution patterns of geosocial media can be further identified, and their spatial anomaly features can be detected. By grouping a set of geographic objects into subsets of similar objects, spatial clustering can identify the aggregation and distribution patterns of the corresponding social media within a certain spatiotemporal range [128]. The spatial clustering methods include partitioning methods (K-means, k-medoids, CLARANS), hierarchical methods (AGNES, DIANA), density-based methods (DBSCAN, DENCLUE), and grid-based methods (STING, CLIQUE). In addition, ST-DBSCAN [129] can be used for spatiotemporal clustering, and partition-and-group [130] techniques can be used for trajectory clustering. Lee and Sumiya [131] first determined the normal state of crowd behavior in an area and then identified the spatial aggregation of tweets based on the number of tweets and the sudden changes in the number of users to enable event detection. By detecting the geographic topic aggregation of tweet streams and their spatiotemporal anomaly characteristics, a real-time local event detector named GeoBurst+ was built to generate event lists automatically and achieves continuous monitoring of data streams [132]. Combined the time and location tags of images, the event collections from Flickr was extracted using suffix tree clustering algorithms [133]. A geographic hierarchical self-organizing map (Geo-H-SOM) model was proposed to analyze the geospatial, temporal and semantic features of tweets and detect their spatiotemporal aggregation patterns [134].

Analyzing the temporal and spatial distribution of security events through social media by incorporating map visualization technology is also a topic of high interest in academic circles. By calculating the tweet ratio and sentiment index, the spatiotemporal distribution of Twitter activity during Hurricane Sandy at scales ranging from global to local was analyzed [17]. Taken the Jemaah Islamiyah bombings in Indonesia as an example, Hastings [40] found that terrorist activities are constrained by space and borders. The spatiotemporal distribution of the terrorist attacks in Iraq between 2004 and 2009 was uncovered using geographic information systems (GIS) method [76]. Related research cases also include the spatiotemporal distribution of terrorist attacks by Euskadi Ta Askatasuna (ETA) in Spain, the geographic distribution of suicide terrorism in Israel [77] and the geographic proximity characteristics of terrorism news on Twitter [42]. Steiger *et al.* [135] systematically summarized the spatiotemporal analyses of Twitter-based social media, indicating that the field is still in the exploration stage.

In summary, location information can be revealed using geotag information or inferred by text feature. In addition, using a combination of text and other features such as social network can enhance the location inference accuracy. Security events can be detected, and their spatiotemporal evolution patterns can be analyzed and tracked using spatiotemporal clustering and anomaly detection. In practice, real-time detection of location information is valuable for geosocial media data using stream computing technology.
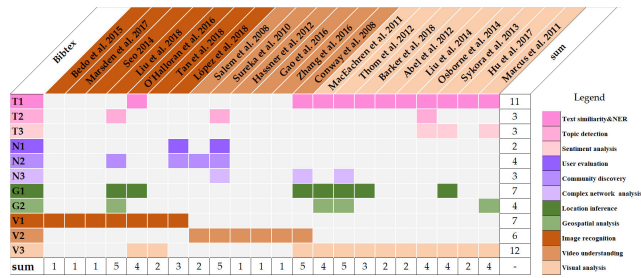
FIGURE 6. 22 selected papers of vision-based analyses for security using geosocial media.

D. IMAGE/VIDEO UNDERSTANDING AND VISUAL ANALYSIS

In social media, image- and video-based analyses and visualization analyses are also becoming increasingly crucial [136]. For example, on Facebook, photo albums, pictures and videos generate more interactions or fan reactions [137] than text posts do, and terrorist organizations such as ISIS widely use images and videos for activities such as terrorist propaganda and recruitment (Figure 6). Traditional forms of image/video understanding used the edge detectors such as Sobel or Canny, key points detectors such as HOG, SIFT, SURF, and ORB, or multi-feature combinations to extract image features, and models such as SVM were used for learning and classification after labeling image/video datasets. Topic models on hashtags from Instagram were proposed to detect protest events using photo location information and spatial autocorrelation indicators [58]. The text and image features from ISIS's online magazine were analyzed to try to understand how they conduct extremist propaganda and how they incite and recruit new members using texts and images [65]. Various reuse patterns of ISIS terror images on media platforms were studied [66]. By extracting local features of the users' images in their profiles and matching them with those of a reference image, the radical users on Twitter can be detected automatically [73]. Images on social media are also an effective propaganda tools during an event or a conflict. The visual themes and structural feature frameworks were uncovered through the contents analysis of 243 images released by the Israel Defense Forces and the Alqassam Brigade on Twitter [55]. In terms of urban security, the FFireDt method was developed for content-based image indexing and classification based on the Flickr platform to identify fire disasters from photo albums on social media [25]. The population density and violent behavior can be detected by deep learning models such as ResnetCrowd [45].

Terrorist organizations have been using online videos for terrorist propaganda and for recruiting and training new members, capitalizing on the Internet to expand worldwide. 50 terrorism videos on YouTube and the 1,443 associated comments were analyzed to identify the characteristics of online users that support such videos [64]. Using content analysis and multimedia coding tools, an exploratory analysis

of 60 Jihadi videos was conducted to identify the video types and group usage patterns [36]. A semiautomated system was built to detect extremist content and users from YouTube and virtual communities and leader users related to extremist videos through user-user and user-video relations [37]. The combination of the image and time features of a video enables the detection of security events, such as crime, violence, and riots through human behavior recognition. Human behavior recognition can be divided into three stages including feature extraction, behavior representation, and classification [138]. An Oriented Violent Flows (OViF) image feature was defined based on the changes in motion amplitude changes in a video to detect violent behaviors in video content [47]. Similar features included the Orientation Histogram of Optical Flow (OHOF) descriptors to distinguish violent behaviors [48].

Visual analysis of multisource and massive social media data can help discover the abnormal characteristics of security events and analyze their temporal and spatial patterns. The SentenTree system can discover of keyword cooccurrence patterns from large-scale social media text collections and aids in the detection of security events [139]. A visual analysis of an event's trends was conducted by used its timeline and revealed its peak to identify important time points during the event [140]. The spatiotemporal visualization of an event can be achieved based on the spatiotemporal information of social media such as SensePlace2 visualization system [11]. By collecting the real-time streaming data from social media, a UK-wide flood situational awareness visualization prototype system was developed to analyze and detect flood events and flood risks [19]. ScatterBlogs, a social media visual analysis system that can perform geographic visualization, was used to extract tweet word cloud to achieve the detection and exploratory analysis of abnormal events on Twitter [15]. Similar visual analysis systems include EMOTIVE for national security [60], WeiboEvents [141], Twitcident [24], and ReDites [38].

In sum, image features have high value for security-related analysis tasks. They can be used to analyze terrorist organizations recruitment and terrorism propaganda, detect violent behaviors, and, when combined, form visual analyses for effectively detecting abnormal features and analyzing the development process of security events. Given the rapid development of computer vision and image analysis technology, it is necessary to further combine the deep learning methods in the image field to improve the accuracy and efficiency of security events detection.

VI. OBSERVATIONS AND DISCUSSIONS

In this literature review, we focused on the geosocial media in the security field. The preceding sections summarized the progress of key technologies related to security events detection and assessing security situations, including natural language processing, social network analysis, location inference and geospatial analysis, and image or video understanding and visual analysis. In this section, we present some observations and discussions.
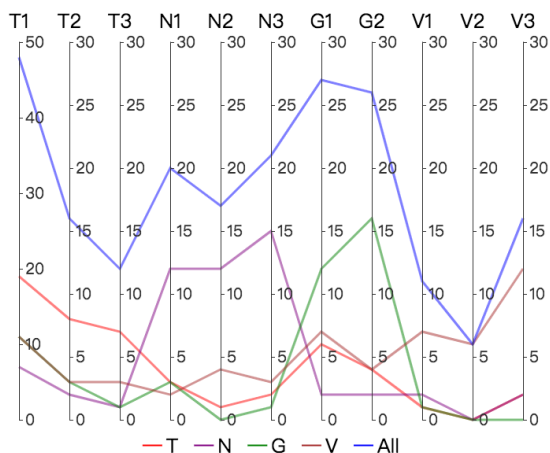
FIGURE 7. The number of cases for different methods.

FIGURE 8. The number of cases for fused-methods.

A. MAIN OBSERVATIONS

Geosocial media-based security research involves two types of analysis tasks. In terms of quantity, security event detection is the focus of current research with 70 cases among all six types of security events. The research on spatial patterns or security situational assessment is relatively rare, mainly focus on spatial analyses of terrorist attacks and risk assessment. Among the 6 types of security events, 16 cases focused on natural disasters, mainly earthquakes or hurricanes; 6 cases addressed man-made disasters involving specific events such as oil spill, urban fires, and nuclear disasters; 22 studied violence related to terrorism events (10 cases); 7 dealt with military-related events during wars or conflicts; 6 were sociopolitical studies that analyzed events such as riots or protests; and 13 other cases were security ideology studies and focused primarily on violent ideologies.

From the perspective of social media platforms, social network sites such as Twitter or Facebook are the primary data sources for various types of research. The majority of studies were based on the data from Twitter (47 papers), while a relatively small number of cases used Facebook data. Some scholars have paid attention to web forums, blogs, and other platforms such as the dark web, Reddit, and online magazines to study terrorist content and features. Related open databases and data sets are also one of the critical data sources, such as the spatial pattern analysis of terrorist activities based on global security databases or the use of GeoNames for security event location. In addition, some studies included image sharing sites such as Flickr or Instagram or video sharing sites such as YouTube, but these involved relatively few cases.

The statistical analyses for these studies was conducted using a variety of technical methods. Figure 7 depicts the statistics of the methods studied in 85 papers. Among the various method types, 19 were studies on text similarity and NER methods, followed by geospatial analysis and complex network analysis (16 articles and 15 articles, respectively). The number of cases applying user evaluation, community detection, location inference, and visual analysis methods
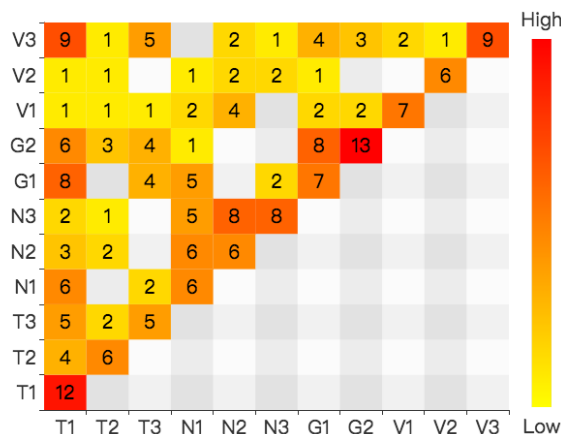
was the same (12 studies). A relatively small number of studies used topic modeling, sentiment analysis, image recognition, and video understanding methods. Overall, more studies focused on text similarity and NER, location inference, spatial analysis, and user evaluation and community detection methods. The most commonly studied topics were text similarity and NER (48 studies). We also analyzed fusion methods (Figure 8) because in many cases of methods using keywords or NER are combined with other methods. For example, there were 9 cases that combined these with visual analysis, 8 cases with location inference, and more than 5 studies that combined geospatial analysis, user evaluation, and sentiment analysis methods. Because of the relevance of the methods, case that fused location inference and geospatial analysis method, as well as topic modeling and community detection were more common (8 studies). In addition, more than 5 papers integrated sentiment analysis and visual analysis methods, user evaluation and topic modeling or community detection. The other fusion methods had relatively few research cases.

B. DISCUSSIONS AND RECOMMENDATIONS
1) COLLABORATIVE SENSING AND DETECTION FOR SECURITY

Currently, many different types of social media platforms exist on which users interact, communicate and share different types of data. The diversity and multisource data features of social media platforms requires collaborative security task awareness and analytical mining. Here, the collaboration involves different aspects, such as data/platform and method. In terms of data/platform, the text/image and social network relations of multiple social media must be collected based on the type of security task, and the text corpus, security database and various types of geographic information are combined, so that multiple security-related clues can be maximally captured to facilitate detection and analysis. For example, multimodal sentiment analysis integrated with multidimensional features, such as text, image, and emojis can overcome some of the limitations of using only text data for

sentiment recognition. In terms of method, multisource data require integrating multiple analysis methods so that the data can be analyzed and cross-tested based on multiple machine learning techniques and pretraining models, thus improving the analysis accuracy. Moreover, to prevent the dissemination of terrorism ideas and terrorist activities, it is necessary to conduct multilanguage text analysis and corpus construction to facilitate more effective event analysis and detection.

2) REAL-TIME SENSING AND DETECTION FOR SECURITY

Security-related analysis tasks require rapid analysis and timely responses, especially for unanticipated security events, which place a higher demand on the real-time nature of security event detection. The related research cases also reflected the value of real-time analysis based on geosocial media data for automatically detecting and monitoring various events [142]. Because of the real-time nature of social media platforms and their characteristics of massive data volume and continuous update characteristics, to achieve rapid security event identification and response, it is necessary to apply big data processing and streaming computing technologies designed to process massive amounts of data in real time. Streaming computing is a computing paradigm for continuous data streams, such as the datasets generated in online transactions and the Internet of Things. These data are borderless and grow constantly, with new records added continuously. The streaming computing paradigm is a process in which data are continuously read from these borderless datasets, processed and results generated. The data on social media platforms are typical streaming data, and streaming computing platforms such as Apache Spark Stream and Flink are needed to continuously collect, process in real-time, and analyze social media data to support rapid responses to security events.

3) DEEP LEARNING FOR NLP AND IMAGE OR VIDEO UNDERSTANDING IN SECURITY

Since the introduction of the deep neural network (DNN), deep learning has made rapid progress in many fields including computer vision and natural language processing [143]. In the field of computer vision, a series of neural network models based on deep convolutional neural network have become widely used in visual computing tasks such as image recognition and classification, object detection, and many others. For example, LeNet-5 was successfully applied to handwritten digit recognition [144]; subsequently, AlexNet [145], VGGNet [146], GoogleNet [147], ResNet [148] and other entries to the ILSVRC and other image classification competitions have continuously improved its classification accuracy. By defining the R-CNN and optimizing it, object positions can be identified, and object detection can be performed [149]–[151]; The YOLO [152] and SSD [153] object detection models were also proposed using the idea of boundary rectangular position regression. Combined with the word vector models such as word2vec or Glove, deep learning models have also made

breakthroughs in the field of natural language processing. In addition to the convolutional network, based on the sequential structure of natural language, the cyclic neural network, RNN was developed. To solve the vanishing gradient problem experienced by RNNs, the LSTM model was defined, which can learn the long-term dependencies in sequence data [154]. Meanwhile, Google proposed the Transformer architecture and built a pretraining model BERT [97], using an attention mechanism, which significantly improved the accuracy of various natural language processing tasks. Because geosocial media involves a large amount of text and image data, the above models and methods can be directly applied to the field of security detection and analysis to improve the analysis results.

4) HETEROGENEOUS DATA ANALYSIS AND EVALUATION FOR SECURITY

The multi-source feature of geosocial media data requires heterogeneous data analysis methods. First, for the same user of social media, multiple attributes and their feature combinations under multiple platforms should be analyzed to better understand their behavior characteristics. Second, auxiliary datasets should be applied to enhance the capabilities of analyzing social media data, for example, in conjunction with city Points of Interest (POI) information to identify spatial semantics of corresponding activities of social media users. Third, social media sensors in cyberspace and internet of things (IoT) sensors in physical space should be integrated under a uniform analytical goal and environment. For example, security awareness and assessment can be performed in conjunction with data in surveillance cameras and social media in a city. Moreover, the evaluation of social media data quality and its reliability needs to be further strengthened to ensure the credibility of the results of security events detection and security assessment.

VII. CONCLUSION

Based on a literature analysis and statistical analysis methods, this survey introduces the recent progress in the use of geosocial media for security-related research. The security-related analysis tasks include two types: security events detection and security situational awareness and assessment. There are six types, including natural disasters, man-made disasters, violent incidents, and military events, sociopolitical events and others security events. We summarized the general process of security-related analysis based on geosocial media, and identified two types of data sets: social media datasets and auxiliary analysis datasets, and discussed the corresponding data acquisition and preprocessing methods. We discussed four major categories of analysis methods and application cases involving natural language processing, social network analysis, geospatial analysis, and image or video understanding. The four types of methods were refined into 11 subtypes, and the studies using the corresponding methods were counted and summarized. The current research focuses on security events detection and involve six types of security events.

The bulk of the studies addressed natural disasters and violence. The main data source is Twitter; however, a few cases exist that acquire data from online forums and photo or video sharing platforms. Natural language processing and geospatial analysis are the primary methods. Some studies integrate multiple methods to perform security analyses. Many cases combine keywords and NER methods with visual analysis and location inference methods. Future work should strengthen the collaborative security analysis approaches using multiple methods, the heterogeneous data analysis and evaluation for security, and combine them with streaming computing technology to conduct real-time security event detection. Deep learning technology should be introduced for natural language processing and images or video understanding in geosocial media. The overall goal is to improve security events detection and security situational awareness or assessment capabilities and accuracy.

ACKNOWLEDGMENT

REFERENCES

[1] M. Kaldor and I. Rangelov, "Introduction: Global security policy in the twenty-first century," in *The Handbook of Global Security Policy*. Malden, MA, USA: Wiley, 2014, pp. 1–8.

[2] UNESCO. (2018). *International Day for Disaster Reduction*. Accessed: Dec. 20, 2018. [Online]. Available: https://en.unesco.org/commemorations/disasterreductionday

[3] A. M. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," *Bus. Horizons*, vol. 53, no. 1, pp. 59–68, 2010.

[4] F. Atefeh and W. Khreich, "A survey of techniques for event detection in Twitter," *Comput. Intell.*, vol. 31, no. 1, pp. 132–164, 2015.

[5] G. Lovink, *Networks Without a Cause: A Critique of Social Media*. Boston, MA, USA: Polity Press, 2011, pp. 1–32.

[6] C. Fuchs, *Social Media: A Critical Introduction*. Salford, U.K.: SAGE, 2017, pp. 1–50.

[7] D. Lazer, D. Brewer, N. Christakis, J. Fowler, and G. King, "Life in the network: The coming age of computational social," *Science*, vol. 323, no. 5915, pp. 721–723, 2009.

[8] CNN. *Two U.S. Airmen Killed in German Airport Shooting*. Accessed: Oct. 9, 2018. [Online]. Available: http://www.cnn.com/2011/WORLD/europe/03/02/germany.shooting/index.html

[9] R. Gupta and H. Brooks, *Using Social Media for Global Security*. Hoboken, NJ, USA: Wiley, 2013, pp. 9–85.

[10] T. Sakaki, M. Okazaki, and Y. Matsuo, "Tweet analysis for real-time event detection and earthquake reporting system development," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 4, pp. 919–931, Apr. 2013.

[11] A. M. Maceachren, A. Jaiswal, A. C. Robinson, and S. Pezanowski, "SensePlace2: Geotwitter analytics support for situational awareness," in *Proc. IEEE Conf. Vis. Anal. Sci. Technol. (VAST)*, Providence, RI, USA, Oct. 2011, pp. 181–190.

[12] A. Hernandez-Suarez, G. Sanchez-Perez, K. Toscano-Medina, H. Perez-Meana, J. Portillo-Portillo, and V. Sanchez. (2018). *Sensing Earthquake Disaster Information: A Named Entity Recognition Approach Using Twitter Collaborative Data*. Accessed: Nov. 20, 2018. [Online]. Available: https://www.preprints.org/manuscript/201808.0269/v1

[13] A. Alfarrajeh, S. Agrawal, S. H. Kim, and C. Shahabi, "Geo-spatial multimedia sentiment analysis in disasters," in *Proc. IEEE Int. Conf. Data Sci. Adv. Anal. (DSAA)*, Tokyo, Japan, Oct. 2017, pp. 193–202.

[14] H. To, S. Agrawal, S. H. Kim, and C. Shahabi, "On identifying disaster-related tweets: Matching-based or learning-based?" in *Proc. IEEE 3rd Int. conf. Multimedia Big Data (BigMM)*, Laguna Hills, CA, USA, Apr. 2017, pp. 330–337.

[15] D. Thom, H. Bosch, S. Koch, M. Wörner, and T. Ertl, "Spatiotemporal anomaly detection through visual analysis of geolocated Twitter messages," in *Proc. IEEE Pacific Vis. Symp. (PacificVis)*, Songdo, South Korea, Feb./Mar. 2012, pp. 41–48.

[16] J. Kim and M. Hastak, "Social network analysis: Characteristics of online social networks after a disaster," *Int. J. Inf. Manage.*, vol. 38, no. 1, pp. 86–96, 2018.

[17] L. Zou, N. S. N. Lam, H. Cai, and Y. Qiang, "Mining Twitter data for improved understanding of disaster resilience," *Ann. Amer. Assoc. Geographers*, vol. 108, no. 5, pp. 1422–1441, 2018.

[18] S. Shen, N. Murzintcev, C. Song, and C. Cheng, "Information retrieval of a disaster event from cross-platform social media," *Inf. Discovery Del.*, vol. 45, no. 4, pp. 220–226, 2017.

[19] J. L. P. Barker and C. J. A. Macleod, "Development of a national-scale real-time Twitter data mining pipeline for social geodata on the potential impacts of flooding on communities," *Environ. Model. Softw.*, vol. 115, pp. 213–227, May 2019.

[20] C. Periñán-Pascual and F. Arcas-Túnez, "Detecting environmentally-related problems on Twitter," *Biosyst. Eng.*, vol. 177, pp. 31–48, Jan. 2019.

[21] K. Banweer, A. Graham, J. Ripberger, N. Cesare, E. O. Nsoesie, and C. Grant, "Multi-stage collaborative filtering for tweet geolocation," in *Proc. 2nd ACM SIGSPATIAL Workshop Recommendations Location-Based Services Social Netw.*, Seattle, WA, USA, Nov. 2018, pp. 1–4.

[22] J. Yin, J. Feng, and Y. Wang, "Social social media and multinational corporations' corporate social responsibility in China: The case of ConocoPhillips oil spill incident," *IEEE Trans. Prof. Commun.*, vol. 58, no. 2, pp. 135–153, Jun. 2015.

[23] Z. Xu, "Crowdsourcing based description of urban emergency events using social media big data," *IEEE Trans. Cloud Comput.*, to be published.

[24] F. Abel, C. Hauff, G.-J. Houben, R. Stronkman, and K. Tao, "Twitcident: Fighting fire with information from social Web streams," in *Proc. 21st Int. Conf. World Wide Web*, Lyon, France, Apr. 2012, pp. 305–308.

[25] M. Bedo, G. Blanco, W. Oliveira, M. Cazzolato, A. Costa, J. Rodrigues, A. Traina, and C. Traina, Jr., "Techniques for effective and efficient fire detection from social media images," 2015, *arXiv:1506.03844*. [Online]. Available: https://arxiv.org/abs/1506.03844

[26] S. Utz, F. Schultz, and S. Glocka, "Crisis communication online: How medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster," *Public Relations Rev.*, vol. 39, no. 1, pp. 40–46, Mar. 2013.

[27] Z. Liu, Z. Wang, S. Chen, Z. Wang, Z. Miao, and X. Yuan, "A platform for collaborative visual analysis on streaming messages," in *Proc. IEEE Conf. Vis. Anal. Sci. Technol. (VAST)*, Paris, France, Nov. 2014, pp. 375–376.

[28] J. Woo, J. Son, and H. Chen, "An SIR model for violent topic diffusion in social media," in *Proc. IEEE Int. Conf. Intell. Secur. Inform. (ISI)*, Beijing, China, Jul. 2011, pp. 15–19.

[29] F. Jordán, "Predicting target selection by terrorists: A network analysis of the 2005 London underground attacks," *Int. J. Crit. Infrastruct. Prot.*, vol. 4, nos. 1–2, pp. 206–214, 2008.

[30] D. Mair, "# Westgate: A case study–how Al-Shabaab used Twitter during an ongoing attack," *Stud. Conf. Terror.*, vol. 40, no. 1, pp. 24–43, 2017.

[31] S. Ghajar-Khosravi, P. Kwantes, N. Derbentseva, and L. Huey, "Quantifying salient concepts discussed in social media content: An analysis of tweets posted by ISIS fangirls," *J. Terrorism Res.*, vol. 7, no. 2, pp. 79–90, 2016.

[32] T. Simon, A. Goldberg, L. Aharonson-Daniel, D. Leykin, and B. Adini, "Twitter in the cross fire—The use of social media in the westgate mall terror attack in Kenya," *PLoS ONE*, vol. 9, no. 8, 2014, Art. no. e104136.

[33] C. Leuprecht and K. Hall, "Why terror networks are dissimilar: How structure relates to function," in *Networks and Network Analysis for Defence and Security*. Cham, Switzerland: Springer, 2014, pp. 83–120.

[34] A. Basu, "Social network analysis: A methodology for studying terrorism," in *Social Networking*. Cham, Switzerland: Springer, 2014, pp. 215–242.

[35] N. Chaurasia and A. Tiwari, "On the use of brokerage approach to discover influencing nodes in terrorist networks," in *Social Networking*. Cham, Switzerland: Springer, 2014, pp. 271–295.

[36] A. Salem, E. Reid, and H. Chen, "Multimedia content coding and analysis: Unraveling the content of Jihadi extremist groups' videos," *Stud. Conflict Terrorism*, vol. 31, no. 7, pp. 605–626, 2008.

[37] A. Sureka, P. Kumaraguru, A. Goyal, and S. Chhabra, "Mining YouTube to discover extremist videos, users and hidden communities," in *Proc. Asia Inf. Retr. Symp.* Berlin, Germany: Springer, 2010, pp. 13–24.

[38] M. Osborne, S. Moran, R. McCreadie, A. Von Lunen, M. D. Sykora, E. Cano, N. Ireson, C. MacDonald, I. Ounis, Y. He, T. Jackson, F. Ciravegna, and A. O'Brien, "Real-time detection, tracking, and monitoring of automatically discovered events in social media," in *Proc. 52nd Annu. Meeting Assoc. Comput. Linguistics*, Baltimore, MD, USA, Jun. 2014, pp. 1–6.

[39] F. Calderoni, "Identifying mafia bosses from meeting attendance," in *Networks and Network Analysis for Defence and Security*. Berlin, Germany: Springer, 2014, pp. 27–48.

[40] J. V. Hastings, "Geography, globalization, and terrorism: The plots of *Jemaah Islamiyah*," *Secur. Stud.*, vol. 17, no. 3, pp. 505–530, 2008.

[41] A. Summa, B. Resch, and M. Strube, "Microblog emotion classification by computing similarity in text, time, and space," in *Proc. Workshop Comput. Modeling People's Opinions, Pers., Emotions Social Media (PEOPLES)*, Osaka, Japan, Dec. 2016, pp. 153–162.

[42] K. H. Kwon, M. Chadha, and K. Pellizzaro, "Proximity and terrorism news in social media: A construal-level theoretical approach to networked framing of terrorism in Twitter," *Mass Commun. Soc.*, vol. 20, no. 6, pp. 869–894, 2017.

[43] A. Elhajj, A. Elsheikh, O. Addam, M. Alzohbi, O. Zarour, A. Aksaç, O. Öztürk, T. Özyer, M. Ridley, and R. Alhajj, "Estimating the importance of terrorists in a terror network," in *Mining Social Networks and Security Informatics*. Berlin, Germany: Springer, 2013, pp. 267–283.

[44] A. E. C. Basave, Y. He, K. Liu, and J. Zhao, "A weakly supervised Bayesian model for violence detection in social media," in *Proc. 6th Int. Joint Conf. Natural Lang. Process.*, Nagoya, Japan, Oct. 2013, pp. 109–117.

[45] M. Marsden, K. McGuinness, S. Little, and N. E. O'Connor, "ResnetCrowd: A residual deep learning architecture for crowd counting, violent behaviour detection and crowd density level classification," in *Proc. 14th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Lecce, Italy, Aug./Sep. 2017, pp. 1–7.

[46] T. Hassner, Y. Itcher, and O. Kliper-Gross, "Violent flows: Real-time detection of violent crowd behavior," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Providence, RI, USA, Jun. 2012, pp. 1–6.

[47] Y. Gao, H. Liu, X. Sun, C. Wang, and Y. Liu, "Violence detection using oriented violent flows," *Image Vis. Comput.*, vol. 48, pp. 37–41, Apr./May 2016.

[48] T. Zhang, Z. Yang, W. Jia, B. Yang, J. Yang, and X. He, "A new method for violence detection in surveillance scenes," *Multimedia Tools Appl.*, vol. 75, no. 12, pp. 7327–7349, 2016.

[49] J. Hammond, "Maps of mayhem: Strategic location and deadly violence in civil war," *J. Peace Res.*, vol. 55, no. 1, pp. 32–46, 2018.

[50] J. Klausen, "Tweeting the *Jihad*: Social media networks of Western foreign fighters in Syria and Iraq," *Stud. Conflict Terrorism*, vol. 38, no. 1, pp. 1–22, 2015.

[51] T. B. Mirani and S. Sasi, "Sentiment analysis of ISIS related tweets using absolute location," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Las Vegas, NV, USA, Dec. 2016, pp. 1140–1145.

[52] M. C. Benigni, K. Joseph, and K. M. Carley, "Online extremism and the communities that sustain it: Detecting the ISIS supporting community on Twitter," *PLoS ONE*, vol. 12, no. 12, 2017, Art. no. e0181405.

[53] T. Zeitzoff, "Using social media to measure conflict dynamics: An application to the 2008–2009 gaza conflict," *J. Conflict Resolution*, vol. 55, no. 6, pp. 938–969, 2011.

[54] T. Zeitzoff, "Does social media influence conflict? Evidence from the 2012 gaza conflict," *J. Conflict Resolution*, vol. 62, no. 1, pp. 29–63, 2018.

[55] H. Seo, "Visual propaganda in the age of social media: An empirical analysis of Twitter images during the 2012 Israeli–Hamas conflict," *Vis. Commun. Quart.*, vol. 21, no. 3, pp. 150–161, 2014.

[56] A. H. Hossny and L. Mitchell, "Event detection in Twitter: A keyword volume approach," 2019, *arXiv:1901.00570*. [Online]. Available: https://arxiv.org/abs/1901.00570

[57] J. Ansah, W. Kang, L. Liu, J. Liu, and J. Li, "SensorTree: Bursty propagation trees as sensors for protest event detection," in *Proc. Int. Conf. Web Inf. Syst. Eng.*, Dubai, United Arab Emirates, Nov. 2018, pp. 281–296.

[58] Z. Liu, X. Zhou, W. Shi, and A. Zhang, "Towards detecting social events by mining geographical patterns with VGI data," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 12, p. 481, 2018.

[59] K. Jung and H. W. Park, "Citizens' social media use and homeland security information policy: Some evidences from Twitter users during the 2013 North Korea nuclear test," *Government Inf. Quart.*, vol. 31, no. 4, pp. 563–573, 2014.

[60] M. D. Sykora, T. W. Jackson, A. O'Brien, and S. Elayan, "National security and social media monitoring: A presentation of the emotive and related systems," in *Proc. Eur. Intell. Secur. Inform. Conf. (EISIC)*, Uppsala, Sweden, Aug. 2013, pp. 172–175.

[61] J. M. Berger and B. Strathearn, "Who matters online: Measuring influence, evaluating content and countering violent extremism in online social networks," in *Proc. Int. Centre Study Radicalisation Political Violence*, 2013, pp. 3–53.

[62] G. L'Huillier, H. Alvarez, S. A. Ríos, and F. Aguilera, "Topic-based social network analysis for virtual communities of interests in the dark Web," *ACM SIGKDD Explor. Newslett.*, vol. 12, no. 2, pp. 66–73, 2011.

[63] I. B. Arpinar, U. Kursuncu, and D. Achilov, "Social media analytics to identify and counter Islamist extremism: Systematic detection, evaluation, and challenging of extremist narratives online," in *Proc. Int. Conf. Collaboration Technol. Syst. (CTS)*. Orlando, FL, USA, Oct./Nov. 2016, pp. 611–612.

[64] M. Conway and L. McInerney, "Jihadi video and auto-radicalisation: Evidence from an exploratory YouTube study," in *Intelligence and Security Informatics*. Berlin, Germany: Springer, 2008, pp. 108–118.

[65] K. L. O'Halloran, S. Tan, P. Wignell, J. A. Bateman, D.-S. Pham, M. Grossman, and A. V. Moere, "Interpreting text and image relations in violent extremist discourse: A mixed methods approach for big data analytics," *Terrorism Political Violence*, vol. 31, no. 3, pp. 454–474, 2016.

[66] S. Tan, K. L. O'Halloran, P. Wignell, K. Chai, and R. Lange, "A multimodal mixed methods approach for examining recontextualisation patterns of violent extremist images in online media," *Discourse, Context Media*, vol. 21, pp. 18–35, Mar. 2018.

[67] A. H. Johnston and G. M. Weiss, "Identifying Sunni extremist propaganda with deep learning," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Honolulu, HI, USA, Nov. /Dec. 2017, pp. 1–6.

[68] B. H. Spitzberg and J. M. Gawron, "Toward online linguistic surveillance of threatening messages," *J. Digit. Forensics, Secur. Law*, vol. 11, no. 3, p. 7, 2016.

[69] T. De Smedt, G. De Pauw, and P. Van Ostaeyen, "Automatic detection of online jihadist hate speech," 2018, *arXiv:1803.04596*. [Online]. Available: https://arxiv.org/abs/1803.04596

[70] J. Qian, M. ElSherief, E. Belding, and W. Y. Wang, "Hierarchical CVAE for fine-grained hate speech classification," 2018, *arXiv:1809.00088*. [Online]. Available: https://arxiv.org/abs/1809.00088

[71] G. K. Pitsilis, H. Ramampiaro, and H. Langseth, "Effective hate-speech detection in Twitter data using recurrent neural networks," *Appl. Intell.*, vol. 48, no. 12, pp. 4730–4742, 2018.

[72] A. Khatua, A. Khatua, and E. Cambria, "A tale of two epidemics: Contextual Word2Vec for classifying Twitter streams during outbreaks," *Inf. Process. Manag.*, vol. 56, no. 1, pp. 247–257, Jan. 2019.

[73] D. López-Sánchez, J. M. Corchado, and A. G. Arrieta, "Dynamic detection of radical profiles in social networks using image feature descriptors and a case-based reasoning methodology," in *Proc. Int. Conf. Case-Based Reasoning*, Stockholm, Sweden, Jul. 2018, pp. 219–232.

[74] C. Bennell and S. Corey, "Geographic profiling of terrorist attacks," in *Criminal Profiling*. Heidelberg, Germany: Springer, 2008, pp. 189–203.

[75] G. LaFree, L. Dugan, M. Xie, and P. Singh, "Spatial and temporal patterns of terrorist attacks by ETA 1970 to 2007," *J. Quant. Criminol.*, vol. 28, no. 1, pp. 7–29, 2012.

[76] R. M. Medina, L. K. Siebeneck, and G. F. Hepner, "A geographic information systems (GIS) analysis of spatiotemporal patterns of terrorist incidents in Iraq 2004–2009," *Stud. Conflict Terrorism*, vol. 34, no. 11, pp. 862–882, 2011.

[77] N. Kliot and I. Charney, "The geography of suicide terrorism in Israel," *GeoJournal*, vol. 66, no. 4, pp. 353–373, 2006.

[78] H. Buhaug, K. S. Gleditsch, H. Holtermann, G. Østby, and A. F. Tollefsen, "It's the local economy, stupid! Geographic wealth dispersion and conflict outbreak location," *J. Conflict Resolution*, vol. 55, no. 5, pp. 814–840, 2011.

[79] D. M. Rusnak, L. W. Kennedy, I. S. Eldivan, and J. M. Caplan, "Analyzing terrorism using spatial analysis techniques: A case study of Turkish cities," in *Evidence-Based Counterterrorism Policy*. New York, NY, USA: Springer, 2012, pp. 167–185.

[80] X. Zhang, M. Jin, J. Fu, M. Hao, C. Yu, and X. Xie, "On the risk assessment of terrorist attacks coupled with multi-source factors," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 9, p. 354, 2018a.

[81] S. M. Radil, C. Flint, and G. E. Tita, "Spatializing social networks: Using social network analysis to investigate geographies of gang rivalry, territoriality, and violence in Los Angeles," *Ann. Assoc. Amer. Geographers*, vol. 100, no. 2, pp. 307–326, 2010.

[82] G. E. Tita and S. M. Radil, "Spatializing the social networks of gangs to explore patterns of violence," *J. Quant. Criminol.*, vol. 27, no. 4, pp. 521–545, 2011.

[83] K. Denecke, *Event-Driven Surveillance: Possibilities and Challenges*. Berlin, Germany: Springer, 2012, pp. 1–51.

[84] M. Atkinson, J. Piskorski, H. Tanev, and V. Zavarella, "On the creation of a security-related event corpus," in *Proc. Events Stories News Workshop*, 2017, pp. 59–65.

[85] Y. Yang, T. Pierce, and J. G. Carbonell, "A study on retrospective and on-line event detection," in *Proc. 21st Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Tampere, Finland, Aug. 1998, pp. 28–36.

[86] M. Cordeiro and J. Gama, "Online social networks event detection: A survey," in *Solving Large Scale Learning Tasks. Challenges and Algorithms*. Cham, Switzerland: Springer, 2016, pp. 1–41.

[87] R. Li, K. H. Lei, R. Khadiwala, and K. C.-C. Chang, "TEDAS: A Twitter-based event detection and analysis system," in *Proc. IEEE 28th Int. Conf. Data Eng. (ICDE)*, Washington, DC, USA, Apr. 2012, pp. 1273–1276.

[88] H. Gu, X. Xie, Q. Lv, Y. Ruan, and L. Shang, "Etree: Effective and efficient event modeling for real-time online social media networks," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol.*, Jan. 2011, pp. 300–307.

[89] K. Bahgat and R. M. Medina, "An overview of geographical perspectives and approaches in terrorism research," *Perspect. Terrorism*, vol. 7, no. 1, pp. 38–72, 2013.

[90] S. L. Cutter, D. B. Richardson, and T. J. Wilbanks, *The Geographical Dimensions of Terrorism*. Oxfordshire, U.K.: Routledge, 2014, pp. 1–65.

[91] R. Pain, "Everyday terrorism: Connecting domestic violence and global terrorism," *Prog. Hum. Geography*, vol. 38, no. 4, pp. 531–550, 2014.

[92] *Global Terrorism Index—Vision of Humanity*, IEP, Sydney, NSW, Australia, 2018.

[93] X. Li and D. Li, "Can night-time light images play a role in evaluating the Syrian Crisis?" *Int. J. Remote Sens.*, vol. 35, no. 18, pp. 6648–6661, 2014.

[94] A. Farzindar and D. Inkpen, *Natural Language Processing for Social Media* (Synthesis Lectures on Human Language Technologies), 2nd ed. San Rafael, CA, USA: Morgan & Claypool, 2017, pp. 1–195.

[95] B. Eriksson, P. Barford, J. Sommers, and R. Nowak, "A learning-based approach for IP geolocation," in *Proc. Int. Conf. Passive Act. Netw. Meas.*, Zürich, Switzerland, Apr. 2010, pp. 171–180.

[96] S. Ribeiro, Jr., and G. L. Pappa, "Strategies for combining Twitter users geo-location methods," *Geoinformatica*, vol. 22, no. 3, pp. 563–587, 2018.

[97] O. Ozdikis, H. Ramampiaro, and K. Nørvåg, "Spatial statistics of term co-occurrences for location prediction of tweets," in *Proc. Eur. Conf. Inf. Retr.*, Grenoble, France, Mar. 2018, pp. 494–506.

[98] T. Miyazaki, A. Rahimi, T. Cohn, and T. Baldwin, "Twitter geolocation using knowledge-based methods," in *Proc. 4th Workshop Noisy User-Generated Text EMNLP Workshop W-NUT*, Brussels, Belgium, Nov. 2018, pp. 7–16.

[99] L. Kong, Z. Liu, and Y. Huang, "SPOT: Locating social media users based on social network context," *Proc. VLDB Endowment*, vol. 7, no. 13, pp. 1681–1684, 2014.

[100] R. Brantingham and A. Hossain, "Crowded: A crowd-sourced perspective of events as they happen," *Proc. SPIE*, vol. 8758, May 2013, Art. no. 87580D.

[101] D. Ahlers, "Assessment of the accuracy of GeoNames gazetteer data," in *Proc. 7th Workshop Geograph. Inf. Retr.*, Orlando, FL, USA, Nov. 2013, pp. 74–81.

[102] M. Haklay and P. Weber, "OpenStreetMap: User-generated street maps," *IEEE Pervasive Comput.*, vol. 7, no. 4, pp. 12–18, Oct. 2008.

[103] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*. [Online]. Available: https://arxiv.org/abs/1810.04805

[104] P. Martí, L. Serrano-Estrada, and A. Nolasco-Cirugeda, "Social media data: Challenges, opportunities and limitations in urban studies," *Comput., Environ. Urban Syst.*, vol. 74, pp. 161–174, Mar. 2019.

[105] A. Ritter, E. Etzioni, and S. Clark, "Open domain event extraction from Twitter," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Beijing, China, Aug. 2012, pp. 1104–1112.

[106] W. H. Gomaa and A. A. Fahmy, "A survey of text similarity approaches," *Int. J. Comput. Appl.*, vol. 68, no. 13, pp. 13–18, Apr. 2013.

[107] A. Ritter, S. Clark, M. Mausam, and O. Etzioni, "Named entity recognition in tweets: An experimental study," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Edinburgh, U.K., Jul. 2011, pp. 1524–1534.

[108] D. Inkpen, J. Liu, A. Farzindar, F. Kazemi, and D. Ghazi, "Location detection and disambiguation from Twitter messages," *J. Intell. Inf. Syst.*, vol. 49, no. 2, pp. 237–253, 2017.

[109] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer, "Neural architectures for named entity recognition," 2016, *arXiv:1603.01360*. [Online]. Available: https://arxiv.org/abs/1603.01360

[110] M. E. Peters, W. Ammar, C. Bhagavatula, and R. Power, "Semi-supervised sequence tagging with bidirectional language models," 2017, *arXiv:1705.00108*. [Online]. Available: https://arxiv.org/abs/1705.00108

[111] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.

[112] D. Ramage, S. Dumais, and D. Liebling, "Characterizing microblogs with topic models," in *Proc. 4th Int. conf. Weblogs Social Media*, Washington, DC, USA, May 2010, pp. 1–16.

[113] W. X. Zhao, J. Jiang, J. Weng, J. He, E.-P. Lim, H. Yan, and X. Li, "Comparing Twitter and traditional media using topic models," in *Proc. Eur. Conf. Inf. Retr.*, Dublin, Ireland, Apr. 2011, pp. 338–349.

[114] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends Inf. Retr.*, vol. 2, nos. 1–2, pp. 1–135, 2008.

[115] B. Liu, "Sentiment analysis and opinion mining," *Synthesis Lect. Hum. Lang. Technol.*, vol. 5, no. 1, pp. 1–167, 2012.

[116] D. Tang, F. Wei, B. Qin, T. Liu, and M. Zhou, "Coooolll: A deep learning system for Twitter sentiment classification," in *Proc. 8th Int. Workshop Semantic Eval. (SemEval)*, Dublin, Ireland, Aug. 2014, pp. 208–212.

[117] A. Severyn and A. Moschitti, "Twitter sentiment analysis with deep convolutional neural networks," in *Proc. 38th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Santiago, Chile, Aug. 2015, pp. 959–962.

[118] F. Riquelme and P. González-Cantergiani, "Measuring user influence on Twitter: A survey," *Inf. Process. Manage.*, vol. 52, no. 5, pp. 949–975, Sep. 2016.

[119] J. Li, A. Ritter, and E. Hovy, "Weakly supervised user profile extraction from Twitter," in *Proc. 52nd Annu. Meeting Assoc. Comput. Linguistics*, Baltimore, MD, USA, vol. 1, Jun. 2014, pp. 165–174.

[120] K. Harrigan, "Geocoding without geotags: A text-based approach for reddit," 2018, *arXiv:1810.03067*. [Online]. Available: https://arxiv.org/abs/1810.03067

[121] X. Kong, J. Zhang, and P. S. Yu, "Inferring anchor links across multiple heterogeneous social networks," in *Proc. 22nd ACM Int. conf. Inf. Knowl. Manage.*, San Francisco, CA, USA, Oct./Nov. 2013, pp. 179–188.

[122] F. Zhou, L. Liu, K. Zhang, G. Trajcevski, J. Wu, and T. Zhong, "DeepLink: A deep learning approach for user identity linkage," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2018, pp. 1313–1321.

[123] I. Pitas, *Graph-Based Social Media Analysis*. London, U.K.: Taylor & Francis, 2015, pp. 10–39.

[124] P. Bhattacharya, S. Ghosh, J. Kulshrestha, M. Mondal, M. B. Zafar, N. Ganguly, and K. P. Gummadi, "Deep Twitter diving: Exploring topical groups in microblogs at scale," in *Proc. 17th ACM Conf. Comput. Supported Cooperat. Work Social Comput.*, Baltimore, MD, USA, Feb. 2014, pp. 197–210.

[125] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, nos. 3–5, pp. 75–174, 2010.

[126] M. E. J. Newman, "Modularity and community structure in networks," *Proc. Nat. Acad. Sci. USA*, vol. 103, no. 23, pp. 8577–8582, 2006.

[127] M. Newman, *Networks*. London, U.K.: Oxford Univ. Press, 2018, pp. 1–56.

[128] J. Han, J.-G. Lee, and M. Kamber, "An overview of clustering methods in geographic data analysis," in *Geographic Data Mining and Knowledge Discovery*, vol. 2. Boca Raton, FL, USA: CRC Press, 2009, pp. 149–170.

[129] D. Birant and A. Kut, "ST-DBSCAN: An algorithm for clustering spatial-temporal data," *Data Knowl. Eng.*, vol. 60, no. 1, pp. 208–221, 2007.

[130] J.-G. Lee, J. Han, and K.-Y. Whang, "Trajectory clustering: A partition-and-group framework," in *Proc. ACM SIGMOD Int. Conf. Manage. Data*, Beijing, China, Jun. 2007, pp. 593–604.

[131] R. Lee and K. Sumiya, "Measuring geographical regularities of crowd behaviors for Twitter-based geo-social event detection," in *Proc. 2nd ACM SIGSPATIAL Int. Workshop Location Based Social Netw.*, San Jose, CA, USA, Nov. 2010, pp. 1–10.

[132] C. Zhang, D. Lei, Q. Yuan, H. Zhuang, L. Kaplan, S. Wang, and J. Han, "GeoBurst+: Effective and real-time local event detection in geo-tagged tweet streams," *ACM Tran. Intell. Sys. Tech.*, vol. 9, no. 3, p. 34, 2018.

[133] M. Ruocco and H. Ramampiaro, "A scalable algorithm for extraction and clustering of event-related pictures," *Multimedia Tools Appl.*, vol. 70, no. 1, pp. 55–88, May 2014.

[134] E. Steiger, B. Resch, and A. Zipf, "Exploration of spatiotemporal and semantic clusters of Twitter data using unsupervised neural networks," *Int. J. Geograph. Inf. Sci.*, vol. 30, no. 9, pp. 1694–1716, 2016.

[135] E. Steiger, J. P. De Albuquerque, and A. Zipf, "An advanced systematic literature review on spatiotemporal analyses of Twitter data," *Trans. GIS*, vol. 19, no. 6, pp. 809–834, 2015.

[136] P. M. Landwehr and K. M. Carley, "Social media in disaster relief," in *Data Mining and Knowledge Discovery for Big Data*. Berlin, Germany: Springer, 2014, pp. 225–257.

[137] HubSpot. *How to Engage Fans on Facebook*. Accessed: Nov. 26, 2018. [Online]. Available: http://cdn1.hubspot.com/hub/53/How-to-Engage-Fans-on-Facebook-04.pdf

[138] H. Wang and C. Schmid, "Action recognition with improved trajectories," in *Proc. IEEE Int. Conf. Comput. Vis.*, Sydney, NSW, Australia, Dec. 2013, pp. 3551–3558.

[139] M. Hu, K. Wongsuphasawat, and J. Stasko, "Visualizing social media content with sententree," *IEEE Trans. Vis. Comput. Graphics*, vol. 23, no. 1, pp. 621–630, Jan. 2017.

[140] A. Marcus, M. S. Bernstein, O. Badar, D. R. Karger, S. Madden, and R. C. Miller, "Twitinfo: Aggregating and visualizing microblogs for event exploration," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, Vancouver, BC, Canada, Mar. 2011, pp. 227–236.

[141] D. Ren, X. Zhang, Z. Wang, J. Li, and X. Yuan, "WeiboEvents: A crowd sourcing weibo visual analytic system," in *Proc. IEEE Pacific Vis. Symp. (PacificVis)*, Yokohama, Japan, Mar. 2014, pp. 330–334.

[142] D. T. Nguyen and J. E. Jung, "Real-time event detection for online behavioral analysis of big social data," *Future Gener. Comput. Syst.*, vol. 66, pp. 137–145, Jan. 2017.

[143] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[144] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[145] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[146] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556

[147] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Boston, MA, USA, Jun. 2015, pp. 1–9.

[148] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.

[149] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 580–587.

[150] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Montreal, QC, Canada, Dec. 2015, pp. 91–99.

[151] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Venice, Italy, Oct. 2017, pp. 2961–2969.

[152] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Las Vegas, NV, USA, Jun./Jul. 2016, pp. 779–788.

[153] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.* Amsterdam, The Netherlands: Springer, Oct. 2016, pp. 21–37.

[154] F. M. Bianchi, *Recurrent Neural Networks for Short-Term Load Forecasting: An Overview and Comparative Analysis*. Berlin, Germany: Springer, 2017, pp. 1–26.

ZHIGANG HAN received the Ph.D. degree from the College of Environment and Planning, Henan University, in 2011. He participated in four projects funded by the National Natural Science Foundation of China. He was a Visiting Scholar with Ryerson University, Canada, from June 2018 to June 2019. He is currently an Associate Professor with the Department of Geographical Information Sciences, College of Environment and Planning, Henan University. His main research interests include related to geographical information sciences and spatial-temporal big data analysis.

SONGNIAN LI received the B.Eng. degree in surveying engineering from the Wuhan Technical University of Surveying and Mapping, Wuhan, China, in 1983, and the Ph.D. degree in geodesy and geomatics engineering from the University of New Brunswick, Fredericton, NB, Canada, in 2002.

From 1983 to 1997, he was a Lecturer, an Associate Professor, and a Visiting Scholar in universities in China and Canada. Since 2001, he has been an Assistant Professor, an Associate Professor, and a Professor with the Department of Civil Engineering, Ryerson University, Canada. His current research interests include geospatial big data, spatio-temporal analysis, moving object and LBS data, geo-collaboration, and geospatial web and web services.

Prof. Li was a recipient of ISPRS U. V. Helava Best Paper Award 2016. He serves as a Council Member of the International Society for Photogrammetry and Remote Sensing (ISPRS), from 2016 to 2020, among other professional societies, was an invited researcher by the Japan Society for the Promotion of Science, in 2009 and 2012.

CAIHUI CUI received the Ph.D. degree from the College of Environment and Planning, Henan University, in 2017. She participated in two projects funded by the National Natural Science Foundation of China. She is currently an Associate Professor with the Department of Regional and Urban Science, College of Environment and Planning, Henan University. Her main research interests include related to spatial analysis and urban geography.

DAOJUN HAN received the Ph.D. degree from the School of Data and Computer Science, Sun Yat-sen University, in 2011. He participated in two projects funded by the National Natural Science Foundation of China. He is currently an Associate Professor with the School of Computer and Information Engineering, Henan University. His main research interests include related to artificial intelligence and formal concept analysis.

HONGQUAN SONG received the Ph.D. degree from the School of Geography, Nanjing Normal University, in 2013. He was a Visiting Scholar with North Carolina State University, USA, from 2015 to 2016. He participated in two projects funded by the National Natural Science Foundation of China. He is currently an Associate Professor with the Department of Geographical Information Sciences, College of Environment and Planning, Henan University, China. His main research interests include related to spatial analysis, video-GIS, air quality modeling, and climate change.

• • •