Title
Spatiotemporal event detection: a review

Permalink
https://escholarship.org/uc/item/1px3q00t

Journal
International Journal of Digital Earth, 13(12)

ISSN
1753-8947

Authors
Yu, Manzhu
Bambacus, Myra
Cervone, Guido
et al.

Publication Date
2020-12-01

DOI
10.1080/17538947.2020.1738569

Peer reviewed
Spatiotemporal event detection: a review

Article in International Journal of Digital Earth · March 2020
DOI: 10.1080/17538947.2020.1738569

13 authors, including:

Manzhu Yu
Pennsylvania State University
38 PUBLICATIONS 306 CITATIONS

Myra Bambacus
NASA
26 PUBLICATIONS 416 CITATIONS

Keith Clarke
University of California, Santa Barbara
255 PUBLICATIONS 11,892 CITATIONS

Qunying Huang
University of Wisconsin–Madison
86 PUBLICATIONS 1,726 CITATIONS

Some of the authors of this publication are also working on these related projects:

- Modeling the environmental susceptibility of landfill sites in California View project
- Spatial computing for geoscience View project
Spatiotemporal event detection: a review

Manzhu Yu\textsuperscript{a}, Myra Bambac\textsuperscript{b}, Guido Cervone\textsuperscript{a}, Keith Clarke \textsuperscript{c}, Daniel Duffy\textsuperscript{b}, Qunying Huang\textsuperscript{d}, Jing Li\textsuperscript{e}, Wenwen Li\textsuperscript{f}, Zhenlong Li\textsuperscript{g}, Qian Liu\textsuperscript{h}, Bernd Resch\textsuperscript{ij}, Jingchao Yang\textsuperscript{h} and Chaowei Yang\textsuperscript{d}

\textsuperscript{a}Department of Geography, Pennsylvania State University, University Park, PA, USA; \textsuperscript{b}NASA Goddard Space Flight Center, NASA Goddard Space Flight Center, Greenbelt, MD, USA; \textsuperscript{c}Department of Geography, University of California-Santa Barbara, Santa Barbara, CA, USA; \textsuperscript{d}Department of Geography, University of Wisconsin-Madison, Madison, WI, USA; \textsuperscript{e}Department of Geography and the Environment, University of Denver, Denver, CO, USA; \textsuperscript{f}School of Geographical Sciences and Urban Planning, Arizona State University, Tempe, AZ, USA; \textsuperscript{g}Department of Geography, University of South Carolina, Columbia, SC, USA; \textsuperscript{h}NSF Spatiotemporal Innovation Center & Dept. of Geography and Geoinformation Science, George Mason University, Fairfax, VA, USA; \textsuperscript{i}Department of Geoinformatics, University of Salzburg, Kapitelgasse, Salzburg, Austria; \textsuperscript{j}Center for Geographic Analysis, Harvard University, Boston, MA, USA

ABSTRACT

The advancements of sensing technologies, including remote sensing, in situ sensing, social sensing, and health sensing, have tremendously improved our capability to observe and record natural and social phenomena, such as natural disasters, presidential elections, and infectious diseases. The observations have provided an unprecedented opportunity to better understand and respond to the spatiotemporal dynamics of the environment, urban settings, health and disease propagation, business decisions, and crisis and crime. Spatiotemporal event detection serves as a gateway to enable a better understanding by detecting events that represent the abnormal status of relevant phenomena. This paper reviews the literature for different sensing capabilities, spatiotemporal event extraction methods, and categories of applications for the detected events. The novelty of this review is to revisit the definition and requirements of event detection and to layout the overall workflow (from sensing and event extraction methods to the operations and decision-supporting processes based on the extracted events) as an agenda for future event detection research. Guidance is presented on the current challenges to this research agenda, and future directions are discussed for conducting spatiotemporal event detection in the era of big data, advanced sensing, and artificial intelligence.

1. Introduction

Spatiotemporal events are occurrences of interest to specific stakeholders and happen at a particular time and location. Examples include natural hazard events, environmental pollution events, health events, urban-related events, business events, and crisis/crime events. Many events bring adverse impacts on the environment and society. For example, natural hazards cost the U.S. $91 billion in 2018 due to 14 hazard events (NCEI 2019); Ebola has inflicted upon the Democratic Republic of the Congo 630 disease cases including 385 deaths since the outbreak in 1976 (ECDC 2019); Natural resource contamination events are costing $4.6 trillion per year in the US (Landrigan et al. 2018). It is crucial to detect the events in real-time and with high accuracy for responding to the challenges to reduce the cost to society, both in lives and in money. For public health events, the Defense
Advanced Research Projects Agency (DARPA) estimated that fatalities can be reduced by a factor of six if the detection time can be advanced by just two days (Neill 2012). However, traditional event extraction methods require reactive human identification or intervention, which significantly reduces the efficiency of event detection (Shi et al. 2018) in addition to the high latency of transmitting sensing data. The latest improvements in sensing capabilities, event extraction methods, and event applications provide new opportunities to facilitate potential timely detections of events. To better understand the emerging technologies and identify future research for enabling such timely detection, we review the latest advances in sensing (Section 2), event extraction methods (Section 3), event applications (Section 4), and discuss the current challenges and future directions (Section 5) in the era of big data (Yang et al. 2019b) and artificial intelligence.

1.1. Definition of a spatiotemporal event

The term ‘event’ has been used a lot in literature with specific examples but seldom defined. An ‘event’ can be defined in three stages progressively (Figure 1) with the first as ‘something that happens at a particular time and place’ (Allan et al. 1998). This definition broadly represents an event in three aspects of what, where, and when, and is accepted in many applications, such as Earth observation, in situ sensing, and environmental monitoring and management (Guillem and Mograss 2005; Metzler, Cai, and Hovy 2012; Atefeh and Khreich 2015). Nevertheless, the definition ignores the process of event including beginnings, duration, ends, and can move across or involve multiple locations, and can be a compound event that has internal events.

The second popular ‘event’ definition is ‘a specific occurrence involving participants’ (Linguistic Data Consortium 2005) by adding the human or agent factor, who, into the aspects. This definition has been largely accepted in computer science and applied fields (Chen, Ji, and Haralick 2009; Elkhilfi and Faiz 2010; Chen et al. 2015), media and communication (McMinn, Moshfeghi, and Jose 2013), and social sciences (Zhou and Chen 2014) for studies on social sensing, wearable sensors, health monitoring and management, urban and business intelligence, and crisis, crime, and social unrest.

The third popular ‘event’ definition is a ‘change of state’ in the monitored measurement (Kopetz 1991) and is shared by a number of approaches in signal processing (Nevatia, Hobbs, and Bolles 2002) and applications (Chen et al. 2005; Zhou and Chen 2014).
2004; Caudal and Nicolas 2005; Hühn 2009; Meira-Machado et al. 2009). For example, a precipitation event occurs when the rate of precipitation suddenly increases (Steele, Goldstein, and Browne 2004). This definition was followed by some publications related to Earth observation, in situ sensing, wearable sensors, environmental monitoring and management, and urban intelligence.

1.2. Requirements of event detection

Spatiotemporal events are complex, dynamic, and context dependent (Ravanbakhsh et al. 2017). Events extracted from various data sources, such as text, imagery, and videos, are related to the characters of the adjacent content (Lee et al. 2015) or disseminated to people and groups affected in a situational awareness fashion. For example, the detection of wildfire needs specific local conditions (closed roads, wind speed and direction, flame height, etc.) for evacuation or routing first-responders (Paschke, Kozlenkov, and Boley 2010). Existing event detection knowledge with situational awareness is usually domain-dependent, based on human annotation or intervention (Mertens et al. 2011) and hard to generalize for cross-domain applications (Kittler et al. 2014).

Spatiotemporal events require monitoring and detection with high accuracy and reliability in content (i.e. whether the event is semantically correct), location, and time to facilitate decision support (Crooks et al. 2013; Liu 2018). For example, disaster relief needs location information as accurate as possible for urgent responses (Singh et al. 2017). Reliability of event detection is critical with trustworthiness and performance consistency for, e.g. protecting drivers from injury and loss of life when detecting a crash. High reliability of energy consumption event detection can contribute to energy conservation and sustainability.

Event detection tasks have different timeliness or real-time requirements ranging from seconds, minutes, days, to a longer time. For example, earthquake warnings are valuable if they precede the earth-shaking that causes damage even a few seconds. Disease outbreak events can be planned for if we can timely index the patients and identify their close contacts during the epidemiologic and clinical investigation (Shmueli and Burkom 2010), e.g. in the coronavirus disease 2019 (COVID-19; WHO 2020). On the other hand, suspicious maritime events, such as arrival at critical points of a vessel trajectory, should be detected within seconds to enable action if necessary (Patroumpas et al. 2017). However, it is challenging to conduct real-time detection due to sensor limitations, data transmission, and semantic ambiguity issues (Tan et al. 2010; Lejeune et al. 2015).

1.3. Spatiotemporal event detection workflow

Instead of considering event detection as the event extraction method alone, we consider spatiotemporal event detection as a collection of processes starting with the sensing, extraction of events of interest, and the applications of actionable events detected (Figure 2).

Event extraction methods focus on understanding the entirety and complexity of an event by identifying (1) the theme or occurrence (i.e. what), (2) the spatial information, such as location or boundary (i.e. where), (3) the temporal information, such as date, time and duration (i.e. when), (4) the participants (i.e. who), and (5) thematic attributes, such as environmental, infrastructural, and social/societal characteristics. Causes and consequences of events may also be identified (i.e. why).

1.4. Review approach

We followed a systematic approach by selecting papers related to ‘event detection’ and ‘spatiotemporal.’ We identified 4,770 publications in Google Scholar and 623 from the Web of Science. We removed duplicates and obtained 4973 papers, and 265 articles were selected based on their keywords most relevant to our topic. Full-text articles were assessed for cutting-edge, centrality to spatiotemporal event detection to reach 132 articles in this review.
2. Sensing

Four major categories of sensing capabilities are summarized and compared for their data sources, characteristics, and major event detection tasks (Table 1).

2.1. Remote sensing

Sensing the Earth remotely has been steadily improved with increased spatial resolution, all-weather imaging, increased spectral resolution, more detailed and frequent viewing, and smaller cost-effective satellites. As a result, events can be detected offline or in real-time and fed into applications such as automated detection and warning systems for planning, policymaking, natural disasters, weather events, crime, and pollution.

Satellites and aboard instruments have been observing optically and thermally the Earth surface for different detections. For example, Landsat 8 (30 m and 16 days revisit) is used for long-term decadal land-use change and fire effect detection (Soulard et al. 2016); Moderate Resolution Imaging Spectroradiometer (MODIS, 1 km and daily revisit) is used for daily active fire detection (Giglio, Schroeder, and Justice 2016); Sentinel 2 (10–50 m and 5 days revisit) is used for detailed agricultural monitoring (Bontemps et al. 2015). The recently launched Geostationary Operational Environmental Satellite (GOES)-R series (5-mins and 0.5 km) provide environmental monitoring, including for hurricanes, lightning strikes, and wildfires for the entire Western Hemisphere.

Active sensors provide clear images with high penetration that enables better capture of surface targets in nearly all-weather and all-time. For example, Synthetic Aperture Radar (SAR) and Light Detection and Ranging (LiDAR) have supported the dynamic monitoring of vegetation and soil moisture (Qiu et al. 2019), ground object change detection (Li et al. 2018), and agricultural cultivation area identification (Tian et al. 2019). LiDAR has also been used in autonomous driving to provide distance measurement (Beer et al. 2018).

Very-high-resolution (VHR) optical images (such as Geoeye, Quickbird, and Worldview) and very-high-resolution SAR images (such as Cosmo-Skymed and TerraSAR-X) offer meter or submeter resolution (Tapete and Cigna 2018) for monitoring detailed surface features that could not be captured before, such as inner water areas, complex land ecosystems, and maritime zones (Nguyen, Kopec, and Netzband 2016; Ibarrola-Ulzurrun et al. 2017).

Hyperspectral and multispectral imaging compose of multiple, 100s, and up to 1000 bands can detect and monitor much more detailed Earth surface targets for, e.g. crop health monitoring, mineral inspection, and fine-scale monitoring of hazardous pollutants (Debba et al. 2005; Adam, Mutanga, and Rugege 2010; Manolakis et al. 2013).
More recently, increasing numbers of cost-effective small satellites are being developed and launched (Nervold et al. 2016; Madry, Martinez, and Laufer 2018) to provide a low-cost platform for various missions, such as establishing constellations for low-cost communication systems and education in satellite development and operation (Gao et al. 2018). Similarly, short-observation time and lightweight hyperspectral imaging sensors have been embedded in unmanned aerial vehicles (UAVs) to provide high spatial resolution (1–50 cm) and flexible operational feasibilities (Zhong et al. 2018).

### 2.2. In situ sensing

Rather than sensing remotely from satellites or aircraft, in situ sensing provides measurements by sensors that are directly located in the environment such as the Earth’s surface, measurement towers, or balloons (Durry and Pouchet 2001; Lawrence and Balsley 2013). In-situ sensors can produce high temporal resolution data instantly and therefore facilitating real-time event detection (Hill, Minsker, and Amir 2007), such as air quality pollution, metabolites in water bodies, the concentration of carbon dioxide (CO2) in living environment.

More recently, wireless sensor networks (WSN) and the Internet of Things (IoT) have provided new opportunities for real time and high-resolution monitoring in a distributed network. The

<table>
<thead>
<tr>
<th>Sensing type</th>
<th>Data sources</th>
<th>Characteristics</th>
<th>Major event detection tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote sensing</td>
<td>Optical, thermal, and radar imagery for Earth monitoring</td>
<td>• Increased spatial resolution&lt;br&gt;• All-weather imaging&lt;br&gt;• Increased spectral resolution&lt;br&gt;• More detailed and frequent viewing&lt;br&gt;• More small satellites for specific purposes</td>
<td>Offline or real-time automated detection and warning systems for planning, policymaking, natural disasters, weather events, crime, and pollution</td>
</tr>
<tr>
<td>In situ sensing</td>
<td>Sensor measurements directly located in the environment, e.g. wireless sensor networks (WSN) and the Internet of Things (IoT)</td>
<td>• High-spatiotemporal-resolution monitoring&lt;br&gt;• Fast information extraction&lt;br&gt;• Sensors are low-cost&lt;br&gt;• Sensor types are heterogeneous&lt;br&gt;• May generate redundant data</td>
<td>(Near) Real-time detection for various types of environmental events, such as air pollution, abnormal metabolites in water bodies, gas leakage, earthquake, and volcano eruption</td>
</tr>
<tr>
<td>Social sensing</td>
<td>Data from location-based devices such as mobile phones and Global Positioning System (GPS) trackers and social media</td>
<td>• Publicly contributed&lt;br&gt;• Individual-based, representing human dynamics and group behavior&lt;br&gt;• Contains noise and biases</td>
<td>Real-time detection for both environmental and social events, such as traffic conditions, environmental pollution, and epidemiology behavior</td>
</tr>
<tr>
<td>Health sensing</td>
<td>Wearable sensors measurements</td>
<td>• Non-invasive, wearable, and mobile&lt;br&gt;• Low power consumption and high energy efficiency&lt;br&gt;• Low signal-to-noise ratio&lt;br&gt;• Limited capability of hardware and software for seamless connectivity and data handling&lt;br&gt;• Discomfort for long-term usages</td>
<td>Real-time detection for:&lt;br&gt;• Abnormal fall events for people with disabilities or the elderly&lt;br&gt;• Behaviors in various kinds of sports activities&lt;br&gt;• Whether a person is in a dangerous situation subject to crime or physical violence&lt;br&gt;• Reveal mental health issues for citizens</td>
</tr>
</tbody>
</table>
flexible distribution of sensors in a WSN facilitates a customized network for various purposes, such as volcano event detection from monitoring sensors (Werner-Allen et al. 2006), spatially and temporally continuous air quality monitoring (Boubrima, Bechkit, and Rivano 2017), and earthquake detection from a distributed acoustic sensor network on regional fiber-optic telecommunication infrastructure (Ajo-Franklin et al. 2019). Compared to the WSN, an IoT system refers to a large collection of sensors used to individually gather and send data through a router to the Internet. In IoT, wireless sensors usually produce a massive volume of high-temporal-resolution data when deployed for fast information extraction and near-real-time applications (Bhuiyan and Wu 2016). This is particularly so when combining high-quality sensing devices with standardized sensing infrastructures to ensure portability and interoperability, paving the way towards the vision of a ‘digital earth’ (Resch, Blaschke, and Mittlboeck 2010; Chen et al. 2014).

Underwater sensor networks and the Internet of Underwater Things (IoUT) have also been proposed to monitor underwater ecology and enhance tsunami warning systems (Kao et al. 2017) with acoustic modems placed in shallow or deep water to detect novel marine animals and extreme events (Cong et al. 2019).

To deal with the heterogeneity, labor-intensive processing and redundancy challenges, studies have been conducted targeting data fusion and standardization (Zhou et al. 2016; Yang et al. 2018), and energy reduction and prolonged lifetime of sensors in the sensor environment.

### 2.3. Social sensing

Social sensing utilizes human sensing capabilities facilitated by location-based devices such as mobile phones and Global Positioning System (GPS) trackers (Liu et al. 2015), and social media platforms (Aggarwal and Abdelzaher 2012). Aggregating individually sensed data could collaboratively represent the group behaviors of a social network (Rosi et al. 2011), such as traffic conditions in an urban area based on GPS sensors, epidemiology behavior changes and interactions from mobile phone co-location (Madan et al. 2010), or the public’s reaction, sentiments and emotions to certain events (Resch et al. 2015a, 2015b).

The emerging social media platforms allow the public to report information ranging from daily life to the latest news intentionally or unintentionally (Xiao, Huang, and Wu 2015; Li et al. 2019). Spatiotemporal events commonly trigger the public to share information in various forms, such as text, voice, images, and videos. The information is especially valuable for event detection processes in a timely fashion, such as disaster or crisis management, when authoritative data are unavailable or not ready (Sakaki, Okazaki, and Matsuo 2010; Zhang et al. 2016; Li, Huang, and Emrich 2019).

Besides social media, Volunteered Geographic Information (VGI) (Goodchild 2007) frameworks is also used to collect data from volunteers and map the impacts or aftermath of natural disaster events through a developed web or mobile application (Longueville et al. 2010). VGI frameworks can also assist military missions or emergency responses to obtain real-time situational awareness of group behaviors and physical events (Jayarajah et al. 2015), such as the Federal Emergency Management Agency (FEMA) app (FEMA 2019).

Except the real-time and informational benefit, social sensing data may contain a significant amount of noise (Huang et al. 2019). Filtering mechanisms have been adopted to extract relevant data and eliminate information that are less trustworthy (Huang et al. 2018; Yang et al. 2019c). Furthermore, the multilingual challenge in the messages published on these platforms also needs to be addressed to better understand and categorize the extracted information (Dang et al. 2014; Lo, Chiong, and Cornforth 2017).

### 2.4. Health sensing

Health sensing uses wearable sensors in direct contact with a human body to monitor the health conditions. These sensors have been used to detect abnormal fall events for people with disabilities or the
elderly (Mubashir, Shao, and Seed 2013), identify behaviors in sports activities, such as snowboarding and football (Kautz, Groh, and Eskofier 2015; Groh, Fleckenstein, and Eskofier 2016), detect whether a person is in a dangerous situation subject to crime or physical violence (López-Cuevas et al. 2017), or reveal mental health issues for citizens (Rodrigues et al. 2015). These wearable sensors leverage the advancement of information and communications technology (ICT) to collect health-related information directly but remotely for doctors and health facilities (Gao et al. 2016; Varatharajan et al. 2018).

Wearable sensors have low power consumption and high energy efficiency. A recent study shows that the quality of data acquired with low-cost (yet certified) wearable sensor technology is comparable to traditional laboratory equipment (Sagl et al. 2019). However, wearable sensors are limited due to the low signal-to-noise ratio, weak data handling, and discomfort for long-term usages (Chen et al. 2016; Majumder, Mondal, and Deen 2017).

3. Event extraction methods

This section summarizes the event extraction methods as rule-based, statistical and probabilistic, image processing, machine learning-based pattern extraction, and simulation with characteristics of each type (Table 2). Note that some methods might be associated with multiple categories, and this list of categories is not exclusive.

3.1. Rule-based methods

Rule-based methods are straightforward and require prior domain knowledge to determine and apply heuristic rules to define the presence of event attributes, such as location, time, and participants. Rule-based methods provide the capability to better capture important information about the topic, location, and time of an event, especially when dealing with noisy datasets. For example, Kilicoglu and Bergler (2009) constructed a dictionary of event triggers to extract possible biological events based on heuristic rules. Ho et al. (2012) designed seven heuristic rules to investigate the geolocation of possible events from online news articles. However, the performance of these rule-based methods varies depending on data quality (Ross, Wei, and Ohno-Machado 2014), rule tuning (Turchin, Gal, and Wasserkrug 2009), and parameter selection (Kaya, Ersoy, and Kamaşak 2009), and the pre-defined rules need to be continuously tuned for better performance (Neill, McFowland, and Zheng 2013; Qin et al. 2013). Therefore, rule-based methods are suitable for situations when the heuristic rules are widely approved or accepted and mainly used to automate the domain knowledge and reduce the labor of manual investigation.

3.2. Spatiotemporal statistical and probabilistic methods

Statistical methods identify various event aspects, such as topic, location, time, participant, and measurement change, based on the noticeable or statistically significant patterns extracted from a sampled sensing dataset. Examples include extracting event topics based on the statistics of volumes of social media keywords within a dynamic period (Chierichetti et al. 2014; Guille and Favre 2015), Extracting event participants, time, and location based on the anomalous social network states that deviate the most from a regular state based on statistical scores, such as empirical p-values that represent pre-defined features within a local spatiotemporal neighborhood (Chen and Neill 2014), measuring changes of an event based on sensor networks, where the statistics of a sliding window on a network can help identify system anomalies (Ge et al. 2015).

Like rule-based methods, statistical methods are also straightforward and generally computationally fast. They usually generate clear patterns and correlations from data and can be easily re-used for sensitivity or variable testing. However, statistical methods are conducted based on a sampled or
generalized dataset from the entire dataset. Whether the sampling procedure (e.g. the spatial and
temporal neighborhood determination) is appropriate for the targeted event extraction task is nor-
manly unknown, and the accuracy of extracted events depends highly on the representativeness of the
sample itself.

Differentiated from statistical methods, probabilistic methods learn the probability distribution of
event occurrence and other related parameters over the entire dataset. For example, Sakaki, Okazaki,
and Matsuo (2010) used Kalman filters to estimate the most probable center locations for event
occurrences, where the data points are assumed to follow the Gaussian distribution. Probabilistic
methods are suitable for extracting events that are dynamic and moving in space. Mousavi et al.
(2013) utilized probabilistic graphical models (PGMs) to monitor the dynamics of a wireless sensor
network. Similarly, Mao, Chen, and Xu (2016) adopted a Markov chain and Bayesian Network (BN)

<table>
<thead>
<tr>
<th>Event extraction method type</th>
<th>Sample methods</th>
<th>Characteristics of method</th>
</tr>
</thead>
</table>
| Rule-based                  | • Heuristics  
• Parameter selection  
• Parameter tuning | • Straightforward and less complex but require prior domain knowledge  
• Performance varies depending on data quality, rule tuning, and parameter selection  
• Suitable for situations when the heuristic rules are widely approved or accepted  
• Mainly used to automate the domain knowledge and reduce the labor of manual investigation |
| Statistical and Probabilistic | • Statistics of social media posts or sensor networks  
• Time series analysis on various time-varying data  
• Bayesian networks  
• Hidden Markov models  
• Gaussian mixture models | • Straightforward and generally computationally fast  
• Generate clear patterns and correlations from data and can be easily re-used for sensitivity or variable testing  
• Accuracy of extracted events depends on the representativeness of the available data set |
| Image processing            | • Pixel-based, object-based or topology-preserving-based methods  
• Object extraction and tracking | • Can explore the dynamics of moving events  
• Lack of concise definitions of objects and events  
• Fuzziness of detection results  
• Hand-coded objects or features relying on the constraints on a limited number of event characteristics |
| Machine learning            | • Spatial and spatiotemporal clustering methods, e.g. k-means, mean-shift clustering, DBSCAN  
• Classification methods, e.g. logistic regression, k-nearest neighbors, deep neural networks, recurrent neural networks, deep convolution neural networks  
• Semantic analysis methods, e.g. Latent Dirichlet Allocation, named entity recognition, self-organizing maps | • Great for image classification, change detection, anomaly detection, and regression-based classification  
• Overfitting issues  
• Supervised methods require sufficient labeled data  
• Unsupervised methods generate results that are hard to explain  
• Difficult to collect datasets describing extreme events that rarely happen  
• Model limited by the available datasets in detecting future events |
| Simulation                   | • Numerical simulation  
• Statistical simulation  
• Agent-based simulation  
• Cellular automata | • Provide alternatives when there is a lack of real-time observation data  
• Provide opportunities to test hypotheses  
• Provide scenarios when experiments or field works are difficult, costly, or almost impossible  
• Highly computing-intensive, especially when there is an increasing demand for higher spatiotemporal resolution forecasts  
• Model uncertainty can be high |
as a spatiotemporal dynamic modeling approach to extract water system events in real-time based on the spatial and temporal dependencies of sensor networks.

However, although probabilistic methods assess and leverage all possible estimations, the estimations are not easily interpretable, as the dependencies between different sensor measurements and between multiple spatiotemporal dimensions are not always transparent. Besides, the estimated results might not be correct if the available datasets are not fully representing the real world.

3.3. Spatiotemporal image processing

Image processing methods have long been applied to extract events from remote sensing imagery and include pixel-based, object-based or topology-preserving-based methods. Pixel-based methods extract events by differencing spectral values at a specific pixel and are less effective in extracting a high level of detail, as well as for dealing with heterogeneous imagery with multi-spatial–temporal-spectral records (Hussain et al. 2013). To overcome this limitation, object-based event extraction methods are developed by considering the spatial context, resulting in higher accuracy in event extraction and reduced misregistration and shadowing effect (Johansen et al. 2010). Topology-preserving image analysis leverages both topological (e.g. adjacency, inclusion, etc.) and non-topological information (e.g. metric distance, angle measure, etc.), whereas both pixel-based and object-based methods ignore topological information (Baraldi and Tiede 2018), e.g. wavelet filters is used to reveal the image texture for the identification of particular objects or spatial features.

Object-based event extraction methods are based on frame differencing, optical flow, and background subtraction (Parekh, Thakore, and Jaliya 2014; Wang and Snoussi 2014; Yu and Yang 2017; Luna et al. 2018), and the extracted objects can be further classified or tracked. For example, Yu and Yang (2017) identified dust storm features by developing a region-growing approach and integrating heuristic and machine learning-based multi-thresholds of dust concentration. Yu, Yang, and Jin (2018) further tracked the movements and seasonal evolvement patterns of dust transport in Northern Africa. Multiple existing methods track the movements of the extracted objects, including point tracking, kernel tracking, and silhouette tracking (Parekh, Thakore, and Jaliya 2014; Mondal, Ghosh, and Ghosh 2016).

Image processing methods bring opportunities to the extraction and tracking of events in space and time. However, due to the lack of concise definitions of objects and events, most existing methods are challenged by the fuzziness of detection results (Karpatne et al. 2018). Besides, most abovementioned methods are still restricted by the hand-coded objects or features, which usually rely on the constraints on a limited number of characteristics (e.g. size, shape, and impact area) of a particular type of object or event.

3.4. Machine learning based pattern extraction

Typical and successful machine learning tasks related to event extraction include image classification, change detection, anomaly detection, and regression-based prediction (Reichstein et al. 2019). There are majorly two categories of machine learning methods: supervised and unsupervised machine learning. The supervised methods use ground truth to learn the relationship between the given sample data and the desired output. The unsupervised methods infer the underlying structure from a set of data points without ground truth. Semi-supervised learning stands in the middle and utilizes a large amount of unlabeled data and a small amount of labeled data to obtain the pattern in general.

Spatial, spatiotemporal, and semantic clustering methods are among the most typical unsupervised ones for event extraction. Spatial clustering methods (e.g. k-means, mean-shift clustering, and clustering using representatives) have been utilized in event extraction application to identify the spatial regions that are impacted by the event (Leichtle et al. 2017). Zhang et al. (2015) utilized a density-based spatial clustering of applications with noise (DBSCAN; Ester et al. 1996) – to
aggregate high-density tweets as a single traffic incident and eliminated low-density regions as outliers. Beyond spatial clustering, spatiotemporal clustering methods have also been developed to extract events integrating both spatial and temporal properties of the sensing data (Sakai and Tamura 2015) for disease outbreaks (Tran-The and Zettsu 2014), public interest and behaviors identification from social sensing, natural disaster impact zone identification (Resch, Usländer, and Havas 2018) and user topic interest and behaviors (Qu et al. 2015).

Classification methods are among the most typical supervised machine learning methods for event extraction. Reynen and Audet (2017) labeled seismic and non-seismic status for each station using seismic attributes (i.e. observed polarization and frequency content) and then trained a logistic regression to classify other stations for seismic event extraction. Tsinganos and Skodras (2018) used a k-nearest neighbor (kNN) classifier to fuse heterogeneous wearable sensor data for the extraction of injury-causing fall events for senior citizens. The kNN classifier classifies a data point based on the majority class of its k nearest neighbors, where k can be determined and tuned by users. Integrating the techniques of deep learning, Liu et al. (2016) applied deep neural networks to classify and extract tropical cyclones and atmospheric rivers from publicly available climate dataset. Chen et al. (2018) developed a change detection method using a Deep Convolution Neural Network (DCNN) to extract the significant surface cover changes before and after the landslide events from high-resolution satellite imagery and aerial images. Reichstein et al. (2019) discussed the integration of spatial and temporal context into deep learning to improve the understanding and prediction of these geoscience processes.

Apart from unsupervised and supervised methods extracting the spatiotemporal and semantic aspects of events, hybrid methods have been developed for real-world applications. Resch et al. (2016) developed a semi-supervised graph-based machine learning to extract semantic information from social media posts to gain a better understanding of urban processes. Urbain (2015) developed a methodology involving named entity recognition (NER) and Bayesian statistics to identify and characterize heart disease risk factor events, and Zhang et al. (2016) presented GeoBurst, a method for real-time local event detection from geo-tagged tweet streams through correlating geo-topics among tweets and comparing them against historical patterns.

One challenge for supervised machine learning methods is that they require sufficient labeled data for the model to learn, but it is challenging to collect datasets describing extreme events that rarely happen. To address this issue, benchmark datasets have been produced, including a solar image benchmarking dataset based on the Solar Dynamics Observatory (SDO; Schuh et al. 2013), surveillance video of 23 types of outdoor activities (Oh et al. 2011), and a change extraction benchmark video dataset with six categories of changes (Goyette et al. 2012). More benchmarking datasets are still in demand to promote a positive and competitive environment for scientific research to produce more effective machine learning methods for event extraction.

### 3.5. Simulation

Simulations, such as numerical simulation, statistical simulation, and agent-based simulation, can provide estimates for potential event when observation data is not available. Numerical simulations are computer programs for a mathematical model representing a physical system. The benefits of numerical simulations is to provide early warning information and enhance disaster mitigation, such as flash flood alert systems (Younis, Anquetin, and Thielen 2008; Alfieri, Thielen, and Pappenberger 2012), earthquake warning systems (Al Amin et al. 2015), and dust storm warning systems (Basart et al. 2012).

Instead of describing systems based on mathematical and physics theories and equations, agent-based simulation and microsimulation utilize simple rules to deduce various possibilities and complexities of behaviors of the system, e.g. for detection of the contamination of a water distribution system and mitigation strategies for potential threats (Zechman 2011), extracting disease-spreading events (Duan et al. 2013) using human contact patterns, and extracting malfunctioning events in a
manufacturing and warehouse management system and optimized the management strategies (Ruiz et al. 2011) considering the resource usage, control of inventory, and providers’ decision making that influence the production level.

However, simulation models can be highly computing-intensive, especially when there is an increasing demand for higher spatiotemporal resolution forecasts (Bauer, Thorpe, and Brunet 2015). Global weather and climate simulations today can achieve a 1 km spatial resolution, based on supercomputing, GPU computing, and artificial intelligence accelerators (e.g. Google Tensor-Flow, Intel Nervana; Fuhrer et al. 2018). Another challenge for simulation is uncertainty that can be partially addressed by assimilating (e.g. 3DVar and 4DVar) observation into simulation (Bennamin et al. 2016). The assimilated data, called reanalysis data, have been widely used for long-term climate studies to extract climate anomalies, including early snowmelt events (Semmens et al. 2013), tropical cyclone intensification events (Martineau and Son 2010), and extreme precipitation (Rivera, Dominguez, and Castro 2014).

4. Spatiotemporal event detection applications

This section reviews the major application types of spatiotemporal event detection. Table 3 shows the mapping among application types, sensing sources, and event extraction methods based on the reviewed literature.

4.1. Environmental monitoring and management

Early detection and forecasting of natural hazards can better prepare us to mitigate the hazardous effects by improving early warning systems with detected spatial, temporal, and magnitude factors. For example, flooding events were detected from street-level tweet incident reports during Hurricane Sandy and an Oklahoma tornado for real-time crisis mapping and management (Middleton, Middleton, and Modafferi 2013a; Middleton et al. 2013b; Yang et al. 2019c). Tsunami events were detected from wireless sensors and IoTs to inform the early warning systems (Poslad et al. 2015). The United Nations has developed the Global Early Warning System for Wildland Fires (Global EWS, http://www.un-spider.org/space-application/user-stories/global_ews) based on the forecast data from the U.S. National Centers for Environmental Prediction (NCEP) Global Forecast System. The U.S. Federal Emergency Management Agency (FEMA) recently released the Integrated Public Alert and Warning System (IPAWS, https://www.fema.gov/integrated-public-alert-warning-system) to provide real-time alerts on various natural hazards around the world and improve coordination of situation awareness and humanitarian aids.

Event detection methods are also applied in monitoring and detecting contamination events of natural resources. For example, water quality conditions are monitored to detect potential pollution using technologies such as hydrodynamic modeling, ICTs, and IoTs (Wang et al. 2015). Similarly, outdoor air pollution has been monitored using IoT sensors (Dutta et al. 2017) and human sensors (social media data) (Jiang et al. 2015). Pollution disasters emitted from liquid or gaseous contaminants have been detected from UAVs (Daniel et al. 2009). Oil spill or oil pollution events have been monitored using synthetic aperture radar (SAR) images (Solberg 2012; Garello and Kerbaol 2017).

4.2. Health monitoring and management

Personal health records can be utilized to support different health needs, such as tracking health status, customizing health plans, and promoting healthy living (Bonander and Gates 2010). Aggregating personal health records can help identify epidemic health issues with significant impacts (Oyana et al. 2017), plan for health service locations (Soares, Dewalle, and Marsh 2017), and conduct health surveillance for intervention and polices (Schinas et al. 2018). Self-monitoring using non-invasive and wearable sensors contribute to cost-effective and easy-to-use strategies for personal healthcare
In addition, information contributed to social media and VGI can also contribute to detecting and monitoring the behaviors of the spatiotemporal patterns of particular epidemic and disease outbreaks (Zadeh et al. 2019).

Health-related event detection can contribute to general disease detection and prevention. Monitoring systems were developed by the Centers for Disease Control and Prevention (CDC) to handle threats from infectious diseases, including the Vaccine Adverse Event Reporting System (VAERS) to monitor vaccine safety (Shimabukuro et al. 2015), and the Global Disease Detection Operations Center to conduct event-based surveillance for diseases worldwide (Christian et al. 2017).

### Table 3. Mapping application types with major sensing sources and event extraction methods.

<table>
<thead>
<tr>
<th>Applications</th>
<th>Sensing sources</th>
<th>Event extraction methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environmental monitoring and</td>
<td>Remote sensing</td>
<td>Rule-based (Giglio, Schroeder, and Justice 2016; Soulard et al. 2016; Qi et al. 2019)</td>
</tr>
<tr>
<td>management</td>
<td></td>
<td>Image processing (Nguyen, Kopec, and Netzband 2016; Ibarrola-Ulurrun et al. 2017; Yu and Yang 2017; Yu, Yang, and Jin 2018)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Machine learning (Hill, Minsker, and Amir 2007; Solberg 2012; Garello and Kerbaol 2017; Chen et al. 2018; Li et al. 2018; Cong et al. 2019; Tian et al. 2019)</td>
</tr>
<tr>
<td>Health monitoring and</td>
<td>Social sensing</td>
<td>Rule-based (Xu et al. 2016)</td>
</tr>
<tr>
<td>management</td>
<td></td>
<td>Machine learning (Sakaki, Okazaki, and Matsuoi 2010; Jiang et al. 2015; Zhang et al. 2016; Resch, Uländer, and Havas 2018; Yu et al. 2019)</td>
</tr>
<tr>
<td></td>
<td>In situ sensing</td>
<td>Statistical and probabilistic (Mousavi et al. 2013; Mao, Chen, and Xu 2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Machine learning (Khanna and Cheema 2013; Schuh et al. 2013; Tran-The and Zettou 2014; Poslad et al. 2015; Dutta et al. 2017)</td>
</tr>
<tr>
<td>Health monitoring and</td>
<td>Social sensing</td>
<td>Rule-based (Kilikoglou and Bergler 2009; Trick 2013; Shimabukuro et al. 2015; Christian et al. 2017)</td>
</tr>
<tr>
<td>management</td>
<td></td>
<td>Statistical and probabilistic (Zadeh et al. 2019)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simulation (Duan et al. 2013)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Machine learning (Missier et al. 2016)</td>
</tr>
<tr>
<td>Health monitoring and</td>
<td>Health sensing</td>
<td>Statistical (Varatharajan et al. 2018)</td>
</tr>
<tr>
<td>management</td>
<td></td>
<td>Machine learning (Kautz, Groh, and Eskofier 2015; Rodrigues et al. 2015; Groh, Fleckenstein, and Eskofier 2016; Oyana et al. 2017; Soares, Dewalle, and Marsh 2017; Schinasi et al. 2018; Birenboim et al. 2019)</td>
</tr>
<tr>
<td>Urban intelligence monitoring</td>
<td>Remote sensing</td>
<td>Rule-based (Ho et al. 2012; Ge et al. 2015; Guille and Favre 2015)</td>
</tr>
<tr>
<td>and management</td>
<td></td>
<td>Statistical and probabilistic (Chierichetti et al. 2014)</td>
</tr>
<tr>
<td>Social sensing</td>
<td></td>
<td>Statistical and probabilistic (Traag et al. 2011)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Image processing (Wang and Snoussi 2014)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Machine learning (Filipponi et al. 2010; Oh et al. 2011; Goyette et al. 2012; Yu, Sun, and Cheng 2012; Jin et al. 2014; Ren et al. 2015; Chang et al. 2016; Cruz-Albarran et al. 2017; Huang et al. 2017; Bok, Kim, and Yoo 2018; Briassouli 2018; Sang, Shi, and Liu 2018; Feller et al. 2019)</td>
</tr>
<tr>
<td>Health sensing</td>
<td></td>
<td>Rule-based (Resch et al. 2015a, 2015b; Osborne and Jones 2017; Werner, Resch, and Loidl 2019)</td>
</tr>
<tr>
<td>Business intelligence</td>
<td>Social sensing</td>
<td>Machine learning (Shroff, Agarwal, and Dey 2011; Verma et al. 2015; Nguyen and Grishman 2016; Lanza-Cruz, Berlanga, and Aramburu 2018)</td>
</tr>
<tr>
<td></td>
<td>In situ sensing</td>
<td>Simulation (Ruiz et al. 2011; Zechman 2011)</td>
</tr>
<tr>
<td>Crisis, crime and social unrest</td>
<td>Remote sensing</td>
<td>Machine learning (Estruch and Alvaro 2012; Rezunovic, Xie, and Grijalva 2013; Akila, Govindasamy, and Sandosh 2016; Ren and Wang 2019)</td>
</tr>
<tr>
<td>monitoring</td>
<td>Social sensing</td>
<td>Machine learning (Wittmer 2015; Tapete and Cigna 2018)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rule-based (Jayarajah et al. 2015)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Machine learning (Li, Chen, and Feng 2012; Artikis et al. 2012; Sakai and Tamura 2015; Ristea et al. 2018)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Machine learning (Li, Chen, and Feng 2012)</td>
</tr>
<tr>
<td></td>
<td>In situ sensing</td>
<td>Rule-based (López-Cuevas et al. 2017)</td>
</tr>
</tbody>
</table>
4.3. Urban intelligence monitoring and management

Urban events are detected for the effectiveness and efficiency of addressing urban issues, such as traffic conditions (e.g. load and collision), noise, energy consumption, safety, and road condition. The increasing population of urban areas leads to increased demands for urban infrastructure and services. The efficient detection of urban events can have a positive impact on the workflow of providing urban services, including health services, energy usage, tourism, and city safety (Pangiotou et al. 2016; Souto and Liebig 2016). By integrating cloud computing and sensing data, the delivery of city services can be enhanced to provide scalable computation and real-time intelligent event detection functionalities (Yu, Sun, and Cheng 2012; Jin et al. 2014). More recently, edge computing and mobile computing have been proposed to extend cloud computing from the sensor end, relieving the computational intensity of a centralized cloud computing resource and enabling real-time processing to support smart cities from the personal terminal to infrastructure hub (Tang et al. 2017; Dautov et al. 2018).

4.4. Business intelligence

Abnormal business-related events are monitored and detected from social sensing to inform business intelligence. Enterprise-related events, such as ‘industrial unrest’ and ‘customer complaints,’ can serve as the input of the enterprise information fusion framework to evaluate the impact of these events and generate real-time alerts (Shroff, Agarwal, and Dey 2011). The uniqueness of business events is that they are future happenings rather than historical happenings (Verma et al. 2015). By monitoring user preferences and claims, companies can formalize the decision-making workflow based on a complex streaming analysis problem for better strategic planning of future events (Lanza-Cruz, Berlanga, and Aramburu 2018). By monitoring business-related events within an enterprise network or manufacturing process, companies can improve the efficiency and agility of the manufacturing processes (Estruch and Álvaro 2012), trace the origin of a network threat (Ren and Wang 2019), identify emerging trends for business transactions (Akila, Govindasamy, and Sandosh 2016), and predict the micro-location patterns of software firms (Kinne and Resch 2018). These detected events can further enable real-time, automatic and systematic reactions based on practical reasoning for business intelligence.

4.5. Crisis, crime, and social unrest monitoring

Social media, surveillance videos, news, and other public records have been monitored to detect crime events, such as shootings (Li et al. 2012), drug deals (Artikis et al. 2012), and civil unrest. The event extraction processes are usually integrated with the streaming data to provide real-time monitoring and reporting. Furthermore, the detected events can also be streamlined with different personnel and agents, such as the police, special security forces, state police departments, and patrol forces. Remote sensing, in combination with surveillance images and videos, has also offered opportunities to detect armed conflicts and terrorist events, especially those in difficult-to-reach and hazardous places (Witmer 2015), and detect the possible factors endangering World Heritage Sites, including armed conflicts, poaching and looting (Tapete and Cigna 2018). These detected events can further inform the early warning system for safety protection of world heritage.

5. Current challenges and future advances

This section reviews the major challenges according to the three steps of spatiotemporal event detection: sensing, extraction methods, and applications (Table 4). Recommended research directions are provided to address these challenges from multiple perspectives.
5.1. Sensing

Sensors have been evolving from pressure sensing to light and sound sensing, infrared, hyper/multi-spectral, ultrasonic, emotion, chemical, and biological sensing. Recent advancements in sensor developments have broadened sensors for commodity: e.g. scanning road for precision and real-time 3D imaging to support autonomous vehicles (Feller et al. 2019); recognizing human gestures with active ultrasonic sensing (Sang, Shi, and Liu 2018); enhancing security with biometric-based fingerprints and facial recognition (Ren et al. 2015); and detecting human emotions with thermal sensing (Cruz-Albarran et al. 2017).

However, it is still challenging to develop more hybrid or multi-sensing elements or systems. The current trend is that sensors are and will be more miniaturized, consume less power, and continuously fuse with other sensors. As an alternative solution, event-driven sensors can remain asleep until an event of interest occurs and might address the issues of energy consumption and short sensor lifetime (Olsson, Bogoslovov, and Gordon 2016).

• Sensor reliability

Sensing results usually have issues of data uncertainty, including inaccurate or missing position information, delayed or out-of-order sensor readings, and the ambiguity of semantics (Liu, Deng, and Li 2017). As reviewed, probabilistic methods are advantageous in dynamic modeling and handling data uncertainty. Appropriate measurements for data correlation are required to better
understand data uncertainty (Koch, Koldehofe, and Rothermel 2008). In addition, extraction methods that are robust to data uncertainty and have low sensitivity need to be developed and experimented with in various sensing environments. For missing data, innovative approaches that estimate or recover missing data need to be further investigated and evaluated.

- **Sensor intelligence**

Sensors are expected not only to capture data of interest but also interpret the data onboard. Onboard processing for satellites, aircraft or UAVs can effectively reduce data volume and data transmission by performing data filtering and decimation on board (Villano, Krieger, and Moreira 2016) or data compression based on hierarchical grid interpolation (Gashnikov and Glumov 2016). These onboard data processing techniques can effectively reduce the latency, enable joint processing from multiple sources, and support further analyses and operations to be conducted (Perez-Neira et al. 2018). However, onboard processing requires the onboard computing device to have high performance and reliability, to use limited resources, and perform in harsh environments due to radiation, temperature, and vibration (George and Wilson 2018). Apart from hardware requirements, the software architecture of the onboard processing for detection needs to be developed and promoted. Sensor intelligence also reflects its capability of learning from the past. Enhanced perception of sensors will help the sensors to self-calibrate and improve the reliability and quality of information towards a smart sensing system.

- **Interoperability between sensors and sensor networks**

Generally speaking, sensor webs (intelligent sensors acting as a whole rather than as a combination of single sensors) have only emerged very recently because of increasingly reliable communication technologies, affordable embedded devices, and the growing importance of sensor data for (near) real-time decision support. One essential driver is the upcoming importance of interoperability, which in this case means that different types of sensors should be able to communicate with each other and produce a common output. Interoperability is a fundamental prerequisite for pervasive sensor webs, involving interoperability at different levels: data structures, measurement transmission, sensor queries, and alerting functionality. Therefore, the Open Geospatial Consortium (OGC) initiated the development of the Sensor Web Enablement (SWE) group of standards (Botts et al. 2006) that serves the whole process chain from data acquisition and transmission to alerting and tasking.

### 5.2. Event extraction method

- **Evaluation metrics are hard to define**

Evaluation metrics of event extraction are hard to define and generalize for different methods and applications. For example, event extraction can be evaluated using quantitative verification, such as the probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI) (Wilks 2011). The classification-based event extraction methods can be evaluated using the classification accuracy metrics, including precision, recall, and $F$-value (Peters and de Albuquerque 2015). The National Institute of Standards and Technology (NIST) defined two evaluation metrics for multimedia event detection $\text{PMiss@TER} = 12.5$ and Normalized Detection Cost (NDC) (NIST 2013). The $\text{PMiss@TER} = 12.5$ refers to the 12.5:1 ratio between the missed detection and false alarm probabilities, whereas the Normalized Detection Cost is a weighted linear combination of these two probabilities. The problem gets more complicated when multiple events are involved with spatial and temporal overlap. Most complex event extraction tasks follow the single event evaluation metrics,
but new metrics need to be developed in cases where contexts are complicated (Temko et al. 2006; Zhuang et al. 2010; Heittola et al. 2011). Evaluation metrics for complex event extraction from heterogeneous sources such as multi-sensor, multi-media, and multi-language are still in demand.

- Cross-domain capability

Existing event extraction methods usually focus on domain-specific events with limits on applicability to cross-domain events. For example, methods extracting landslides are not necessarily capable of extracting extreme precipitation. Besides, features constructed by manual identifications or pre-defined rules are incapable of producing domain-invariant elements for cross-domain event extraction (Xu et al. 2016). Depending on the capability of different event extraction methods, methods may have the potential to be adapted for a certain number of domains or events (Nguyen and Grishman 2016; Yu et al. 2019). Literature has also discussed the adaptability of their methods to other domains (Sagl et al. 2012), though these methods are rarely developed for cross-domain purposes.

- Limited training/benchmark data

We also observe the negative impact of limited training or benchmark data on the enhancements of event extraction methods. It is extremely difficult to overcome the issue due to limited training data because the learned rules or models might contain substantial biases (Chang et al. 2016). For example, adequate validation accuracy cannot be achieved due to limited training data (Ahmad et al. 2016), and a deep learning architecture easily leads to overfitting of the limited training data (Nie et al. 2017). For reducing the impact of limited training data, data augmentation techniques, e.g. flipping, distorting, or cropping training images, have been applied to increase the size and diversity of training data (Li and Hsu 2018). However, these data augmentation techniques do not fundamentally address the challenge of the limited training data issue. More open benchmark data should be advocated and developed with the purpose of research sharing and competition.

- Machine and Deep Learning

Many existing event extraction methods learn from past events using the statistics and probabilities of event occurrences and properties. However, the contexts of events, even of the same type, are changing rapidly over space and time, making it difficult for existing event extraction methods to produce accurate results for future events (Yu et al. 2019). Besides, existing methods may fail when dealing with an unknown type of event, especially when there is a lack of training data. Machine and deep learning have expanded event detection systems to identify unknown events by studying the potential triggers of unknown events (Atefeh and Khreich 2015), exploring the abnormal spatiotemporal patterns in the data sources (Briassouli 2018), or transfer the learning of existing human-constructed event schemas to detect unseen events (Huang et al. 2017). Future methods and applications need to have the capability of life-long learning where a system can constantly identify new, unknown events and incrementally add these events to the knowledge base.

The interactivity and explainability of the machine and deep learning methods are increasingly crucial. Machine and deep learning-based event extraction methods with interactivity can enhance human-machine communication and correct errors in the early stages, whereas the methods with explainability can improve the trust of users for the detected events. Furthermore, due to the urgent need for real-time response and decision making, new self-active learning or even unsupervised learning methods should be developed for event detection with limited training datasets and as little human annotation effort as possible.
5.3. Application

- Sensing data preprocessing for actionability of event detection applications

The exponential increase of sensing data volume has required high-performance computational resources for scalable event detection (Yang et al. 2019b). Computational platforms have been developed to meet this requirement, such as Google Earth Engine (Gorelick et al. 2017), Amazon Web Services (AWS, https://registry.opendata.aws/), and European Space Agency (ESA) Thematic Exploitation Platform (TEP, https://tep.eo.esa.int/), providing online access to archived EO data (e.g. Landsat, Sentinel, and MODIS). However, functionalities of handling the big spatiotemporal data are not readily available regarding query optimization, preprocessing, and streaming information extraction from the massive spatiotemporal data (Yang et al. 2017; Li et al. 2017a, 2017b, 2019; Yang et al. 2019a).

Any single event extraction problem may consider a variety of diverse data sources with different data types and formats. Therefore, fusing data from multiple sources, such as multi-sensor, multimedia, and multi-language, is crucial through aggregating, converting, and reformatting into a uniform structure (Zhang 2010). Data fusion techniques remain challenging for multi-source data fusion within different resolutions or internal structures, data imperfection, ambiguities and inconsistency, semantic gaps, and information conflict (Khaleghi et al. 2013; Alam et al. 2017).

- Transfer detection results to operation and decision support

For event processing in real-time, it is essential to integrate event detection and the results of the detection into operational systems. Architectures have been proposed to better integrate event detection in various kinds of applications, such as subway emergency operations (Filippone et al. 2010), and processes can be reused for efficient event detection and decision-supporting purposes (Bok, Kim, and Yoo 2018). By archiving previous events, event extraction methods can learn the previous conditions for similar events to occur and the following procedures for relevant operational agents to take action (Kezunovic, Xie, and Grijalva 2013). The challenges of transferring event detection results into operational systems are majorly due to the inability to detect events with high accuracy in real-time. Exploring increasingly available types of data sources might improve the spatial and temporal resolution and coverage to facilitate higher accuracy of event detection. Further investigations are still needed through addressing the challenges of sensing data and event extraction methods.

To make informed decisions following detected events, decision-makers will need timely and accurate information about events’ possible consequences, public perceptions of those consequences, and any available negative impact mitigation options. Therefore, event detection is only a part of the decision-making processes. However, the challenge of transferring detection results to decision-makers lies in the capability of event detection community to explain and visualize the detection uncertainty (Artikis et al. 2012; Khanna and Cheema 2013). Using the information provided by event detection, risk managers need to quantify and characterize the detected events along with the scientific uncertainties (Li, Chen, and Feng 2012; Trick 2013). Decision-makers decide based on the analysis of risk managers whether a perceived risk warrants action. The scientific uncertainty, by itself, does not directly impact the decision-making.

6. Conclusion

As an overarching objective of almost all remote sensing and many other sensing operations, spatiotemporal event detection is a critical foundation for the advancement of sensing technologies. We analyzed recent advancements in spatiotemporal event detection by revisiting the definition of
spatiotemporal events and the methods of detection. We aim to draw the reader’s attention to the scientific and technological advancements of sensing technologies, event extraction methods, and applications for better accuracy, effectiveness, and efficiency. A research objective for spatiotemporal event detection is laid out, including analyzing varied sensing data, leveraging available technologies to extract spatiotemporal events, and mitigating the potential negative impacts of these events. The main challenges remain in effective handling of sensing data, developing superior extraction methods associated with specific domains and applications, and utilizing detected events for operational processes.

Overall, challenges exist in: (a) sensing technologies for sensor design, sensor reliability, and sensor intelligence; (b) event extraction methods in hard-to-define evaluation metrics, limited cross-domain capability, and limited training and benchmarking datasets; (c) ensuring the actionability of detected events, and how to integrate or transfer event information into operational and decision-support systems while considering uncertainty propagation, timeliness, and other factors. Future work should address these challenges by developing novel real-time event detection methods. In addition, with the rapid increase of the available spatiotemporal data sources, future research should better integrate these data for comprehensive and multi-dimensional event detection to benefit both the scientific understanding of events and the operational processes of event-related decision making.

Acknowledgements

We thank Drs. Changlin Wang, Zhen Liu, and the anonymous reviewers for their constructive comments and suggestions. The research is supported by NSF (CNS 1841520 and ACI 1835507), NASA (80NSSC19P2033), and the NSF Spatiotemporal I/UCRC IAB members. This community paper is one of a set of efforts to setup the foundation for future spatiotemporal studies/sciences. Yang, Yu and Bambacus initialized the concept for this paper. Yu and Yang designed the paper structure. Yu wrote the paper with assistant from other authors. Yang revised the paper with other authors. Other authors contributed equally to the paper for idea, writeup, and revision. They are listed in alphabetical order based on last name.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The research is supported by NSF [CNS 1841520 and ACI 1835507], NASA Goddard [80NSSC19P2033], and the NSF Spatiotemporal I/UCRC IAB members.

ORCID

Keith Clarke http://orcid.org/0000-0001-5805-6056
Jing Li http://orcid.org/0000-0002-9544-8240
Wenwen Li http://orcid.org/0000-0003-2237-9499
Jingchao Yang http://orcid.org/0000-0001-7071-7291
Chaowei Yang http://orcid.org/0000-0001-7768-4066

References


Hühn, P. 2009.


