Big Data and Emergency Management: Concepts, Methodologies, and Applications

Xuan Song, Haoran Zhang, Rajendra Akerkar, Huawei Huang, Song Guo, Lei Zhong, Yusheng Ji, Andreas L. Opdahl, Hemant Purohit, André Skupin, Akshay Pottathil, Aron Culotta

Abstract—Recent decades have seen a significant increase in the frequency, intensity, and impact of natural disasters and other emergencies, forcing the governments around the world to make emergency response and disaster management national priorities. The growth of extremely large and complex datasets —commonly referred to as *big data* —and various advances in information and communications technology and computing now support more effective approaches to humanitarian relief, logistical coordination, overall disaster management, and long-term recovery in connection with natural disasters and emergency events. Leveraging big data and technological advances for emergency management has attracted considerable attention in the research community. However, the desired merging of *big data and emergency management* (BDEM) requires coordinated efforts to align and define interdisciplinary terminologies and methodologies. To date, the key concepts and technologies in this emerging research area have not been coherently discussed in a sufficiently broad and multidisciplinary manner. In this article, an international team presents an overview of the BDEM domain, highlighting a general framework and discussing key challenges from several perspectives. We introduce and summarize typical technologies and applications, organized into the six broad categories of remote sensing, resilient communication networks, mobile communication networks, human mobility and urban sensing, social network analysis, and knowledge graphs. Finally, we outline several directions of future research.

Index Terms—Disaster Informatics, Urban Computing, Smart City, Emergency Management.

1 Introduction

Hurricanes, earthquakes, floods, terrorist acts, and catastrophic infrastructure failures cause immense physical destruction and catastrophic loss of life and property around the world. In recent decades, the frequency, intensity, and impact of disastrous events have increased significantly. Faced with such events, governments have recognized that emergency response and disaster management are major concerns requiring concerted efforts in collaboration with business and academia. Contemporary information and communications technology and the emergence of extremely large and complex datasets (i.e., *big data*) have made it possible to employ advanced techniques to reveal patterns, trends, and associations that enhance situational awareness and decision-making in emergency scenarios. Data-driven emergency response and management have been successfully applied in many recent events, including large-scale natural disasters.

For example, real-time mapping of road conditions during an emergency, from congestion to blockages, allows intelligent support for emergency vehicle navigation, transportation of rescue teams, distribution of supplies, and dispatching of volunteers. During the 2011 East Japan Earthquake, the road network was severed numerous times across a very wide area. In response, actual traffic data collected via GPS sensors from moving vehicles was used to derive a real-time road map. This so-called "probe information service," based on collecting GPS sensor data via the car navigation systems of subscribed vehicles, was originally intended for daily monitoring of road traffic. In 2006, the use of this service for purposes of disaster response was studied by ITS Japan (NPO) and Honda Motor Co. During the 2011 earthquake, the probe information service was deployed to aggregate daily traffic data from multiple private companies, including Honda, Pioneer, Toyota, and Nissan. This allowed the generation of high-fidelity road passage maps that greatly aided disaster recovery activities [1].

Numerous other examples have led to the use of advanced computing and big data for emergency management attracting considerable attention from the research community. However, the delineation of a coherent research agenda remains elusive.

1.1 Big Data and Emergency Management

A comprehensive approach to emergency management is a difficult endeavor, involving access to diverse types of information through heterogeneous channels that can respond to dynamic updates over time. Indeed, *big data and emergency management* (BDEM), as an emerging research area, springs from fundamentally different intellectual lineages and societal contexts. There is innate tension between the technological capabilities and well-meaning ambitions

- *X. Song (E-mail: songxuan@csis.u-tokyo.ac.jp) are with Southern University of Science and Technology (SUSTech), China and The University of Tokyo, Japan. H. Zhang are with The University of Tokyo, Japan. R. Akerkar is with Western Norway Research Institute, Norway. H. Huang and S. Guo are with The Hong Kong Polytechnic University, Hong Kong. L. Zhong and Y. Ji are with National Institute of Informatics, Japan. A. Opdahl is with University of Bergen, Norway. H. Purohit is with George Mason University, USA. A. Skupin and A. Pottathil are with San Diego State University, USA. A. Culotta is with Ilinois Institute of Technology, USA.*

Manuscript received ; revised

of big data (BD) experts and the real-world needs and constraints of emergency management (EM) practitioners.

Over the past five years, several survey articles have summarized existing terminology and methodologies [2], [3], [4], [5] in data-driven disaster management. However, these summaries typically only cover state-of-the-art or key technologies in one or two specific research areas, such as data mining, machine learning, or communication networks. However, BDEM involves cross-domain terminology and methodologies and should be carried forward by multidisciplinary and international efforts. *To this end, we formed a multidisciplinary team of 13 experts from nine institutions with backgrounds ranging from computer science to communications engineering, remote sensing, cartography, and GIS.* The main focus of this paper is to present an overview of major BDEM concepts, introduce an overarching framework and key data sources, and discuss the main challenges from an interdisciplinary perspective. We introduce and organize the core BDEM technologies and applications into six categories, namely remote sensing technologies, resilient communication networks, mobile communication networks, human mobility modeling and urban sensing, online social network analysis, and knowledge graphs. Finally, we provide our key findings, present an outlook on the future of the BDEM domain, and suggest several future research topics.

The remainder of this article is structured as follows. Section 2 provides an overview of the concept of *big data and emergency management* and discusses the key challenges in the field. Section 3 covers remote sensing technologies for EM. Sections 4 and 5 describe resilient communication networks and mobile communication networks for EM. Section 6 then introduces human mobility modeling and urban sensing for EM. Sections 7 and 8 introduce online social network analysis and knowledge graphs for EM. Finally, Section 9 reflects on the state of the domain and delineates ongoing research challenges and topics for consideration by the community.

2 Overview of the Concept

2.1 Overview of the General Concept

Figure 1 depicts the general concept of BDEM. The main purpose of BDEM is to explore effective methodologies and sustainable solutions to screen the critical and heterogeneous information, detect and predict relief demands, help managers to make effective decisions, then, integrate them to construct a real-time knowledge system.

During a crisis, everybody involved - the public, the media, the government, emergency services, relief organizations, and others- can contribute towards prompt situational awareness. Social network information, population location information, and sensor information represent the main channels through which people collectively build awareness, with the advantages of being distributed, far-reaching, and instantaneous. And as the disaster evolves, the quantity and quality of the information will also grow as illustrated by the timeline. To effectively collect, transmit, extract, analyze, conserve, and utilize the information, three key technical parts are essentially needed which are infrastructural technology (resilient communication networks, mobile communication networks), analysis technology (remote sensing technologies, human mobility modeling and urban sensing, online social network analytic), and superstructural technology (knowledge graphs). We describe the overall framework in the following lines.

In large-scale emergencies, data-driven emergency response is usually quite challenging under the state of emergency as the underlying communication networks are also disrupted. Thus, building resilient communication networks for data collection and dissemination is essential for effective emergency response. Besides, the emergency communication networks (ECNs) need to be immediately established to respond to post-disaster operations. For example, the ECN Center can collect messages from disaster areas and notify the victims of the required actions for disaster-relief. When a group of smart-phones connect to the Internet through cellular or Wi-Fi networks, or when they connect together under a certain topology using the built-in Bluetooth technology, a sensor network is actually constructed and can be exploited to organize ECNs for disaster-relief tasks.

While the natural disasters or emergency events occur, satellite remote sensing systems offer valuable observation data (e.g. satellite imagery) for constant monitoring of atmospheric and surface patterns linked to natural emergencies. Furthermore, satellite communication is essential for operative emergency response, especially in position location, alerting, data collection, and harmonizing relief procedures. Meanwhile, based on human mobility and urban sensing data, it is critical for modeling and predicting human behavior and their travel routes in order to design appropriate transportation scheduling, humanitarian aid, and emergency management. At the same time, social media messages during emergencies and disasters can be analyzed to model and predict human evacuation behaviors and collect relevant information such as caution-advice and damage reports, request and offer to help as well as emotional support for the affected community. After integrating these results, the disaster development can be emulated. Considering the nature of the data is highly heterogeneous, and from multiple sources with varying levels of quality and correctness, a collaboration mechanism will be established within each development step to cleanse the chaotic information.

Finally, based on the constructed resilient communication networks and smart-phone based ECNs, the big and heterogeneous data source and analysis result will be collected that include automated and human sensor data, mobility and communication data, online social network data and open and government data. Knowledge graphs and the semantic technologies they build on offer the unforeseeable ways to make these data be rapidly interpretable and available for complex and querying, processing, and reasoning.

2.2 Key Challenges

As an emerging field, BDEM faces tremendous challenges as follows:

Sensing technologies and data acquisition: Following natural or man-made disasters, effectively sensing and monitoring the event status while unobtrusively and continually

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2020.2972871, IEEE Transactions on Big Data

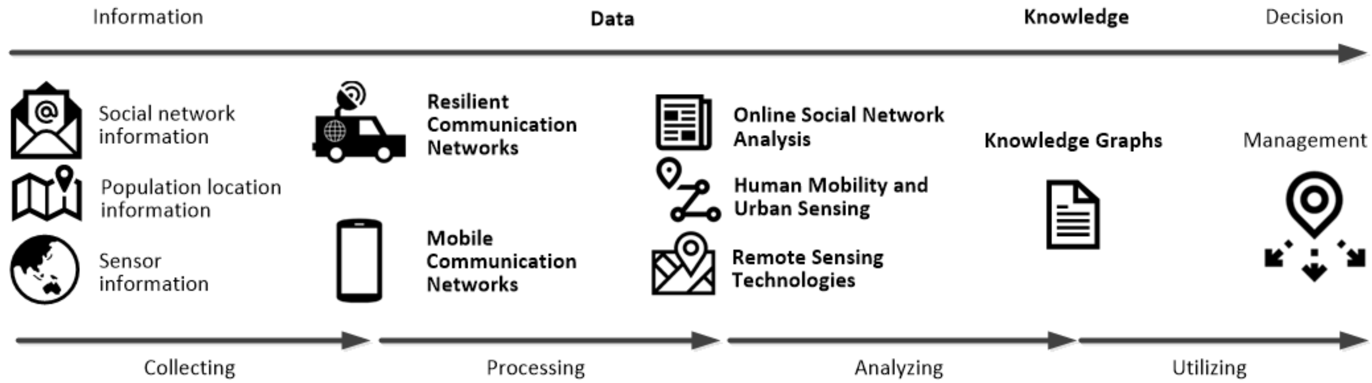


Fig. 1: Overview of the General Concept.

collecting observation data will become a key problem for emergency management teams. Nowadays, users have to evaluate the nature of the emergency and select the appropriate sensor for the job. One sensor might be appropriate in terms of spectral resolution, but offer limited temporal resolution and become unusable under cloud coverage.

Disasters can potentially destroy the power supply infrastructure and cause massive outages. Thus, maintaining the live status of large numbers of sensing devices with a limited power budget becomes an intractable challenge. In the era of big data, human sensing is an emerging approach for tackling this challenge. However, this modality faces challenges in terms of privacy, loosely controlled and nonuniform distributed sensors, and unstructured, implicit, and noisy data, as discussed in [6].

Computing with big and heterogeneous data: Though there is an enormous amount of big data (such as social media data [7], human mobile GPS data [8], and satellite image data [9]) involved in emergency management, tackling various issues related to the volume, variety, velocity, variability, validity, volatility, vulnerability, and veracity of these data are vital in utilizing all available information. There is a need to investigate new techniques for integrating data from distinct resolutions and spatiotemporal scales with different emergencies.

Additionally, how to effectively utilize a deluge of spatial information under the development of Internet of Things (IoT) is a key challenge. [10]. In the real world, time-critical data from sensors and social media increases rapidly in any given emergency event. Therefore, it is essential to develop new analytical tools and techniques, which will create a new level of abstraction and provide novel insights into the patterns of the emergency event.

Human behavior understanding and modeling: During disaster scenarios in urban districts (for example, a large fire in a tall building or an explosion in a crowded area), people easily become panicked, which can result in casualties. Thus, providing efficient evacuation guidance to the crowd is of great importance. This can be realized by understanding and modeling the human mobility of the crowd following a disaster.

Obviously, the modeling, perception, and accurate forecasting of human behavior during an emergency will play a critical role in efficient emergency management. However, human behavior and movement patterns usually have excessive degrees of freedom and inconsistencies, and are impacted by many factors (e.g., weather, urban structure, event type, traffic conditions). Thus, human mobility and behavior are very difficult to model and predict.

Big data for relief decision making: As the "terminal" of the emergency management system, emergency decision making (EDM) is receiving more and more attention. Mathematical models are the most common way to conceptualize the optimization problem in relief. When facing emergencies, different mathematical models should be established according to different requirements. There are three main categories, including evacuation oriented optimized scheduling model, resource allocation optimization model, and logistics optimization model.

Usually, the proposed mathematical models are NP-hard when considering many practical constraints and the results cannot be obtained in polynomial time. When taking the "deluge" of information of emergency situation into consideration, solving the mathematical models are time-consuming. How to coupling the above preanalysis techniques to realize the real-time optimization for relief decision making will be challenging works.

3 Resilient Communication Networks for Emergency Management

Data-driven emergency management usually requires distributed data to be collected in real time due to the distributional and exceptional nature of an emergency. Unfortunately, this is quite challenging during an emergency, as the communication networks are usually disrupted. In large-scale emergencies, many resources, from societal infrastructure to computing and communication resources, may become unavailable or disrupted over a large geographical area and for a long period of time, depending on the type, scale, and degree of the emergency. Up to 1.9 million fixed communication lines and $29,000$ base stations (BSs) were damaged by the 2011 Great East Japan Earthquake [11]. In spite of this, the use of computing and communication resource during an emergency is important, and there may be a much higher demand for resources in both the stricken area and the wider community. For example, during the 2011 Great East Japan Earthquake, fixed-line phone traffic was 4–9 times the normal level (NTT East), and mobile

Fig. 2: Resilient communication networks for Big Data collection and information dissemination in emergency management.

phone calls increased by a factor of 50–60 (NTT DoCoMo). Under such conditions, carriers had to block 80–90% of fixed phone calls and 70–95% of mobile phone calls [12]. Thus, the physical damage and/or congestion of communication networks makes real-time information collection and dissemination for emergency management quite challenging.

Building resilient communication networks for data collection and information dissemination in emergency management is essential. The data flows during an emergency are depicted in Fig. 2, where various kinds of data are collected from the emergency sites (on the left) through the communication networks, and the emergency response decisions based on BD analytics, such as evacuation guidance and rescue instructions, are disseminated to their destinations (on the right). Useful information, such as road conditions, availability of medical help, and survivor locations, can be derived from the big data and external static data, such as map information. To disseminate information that assists survivors, rescue activities, and other kinds of disaster response activities, a working network is a necessity. Hence, communication methods provided by various network facilities play a key role in the emergency response. Mechanisms for discriminating among the raw data collected from sources such as environmental sensors and mobile phones, efficiently distributing information, refining the information, and reforming computing and communication resources based on this information are the keys to ensuring the smooth flow of information for emergency response.

3.1 Key Technologies and Methodologies

During an emergency event such as a natural disaster, we need to know and inform people about safe locations, food sources, resources such as hospitals and gas stations, and the road conditions. To collect and disseminate such information, we need a functional communication network. However, many network elements (nodes and links) may have been destroyed during the disaster. Therefore, a communication network that can survive a tremendous range of failures is instrumental in any data-driven emergency management system. Given the intrinsic difference between the backbone and access points of communication networks, we consider two different approaches for each: fast resilient mechanisms for backbone networks and on-site networking for wireless access networks.

Resilient communication networks can swiftly recover when the communication infrastructure is partially damaged during an emergency. In the backbone part, this is possible because the networks are spread over a wide geographical area so that, even when a major disaster occurs, part of the backbone network resources will still be available. As access networks are usually located over small geographical areas close to users, the whole network may be destroyed by an emergency. Hence, constructing a new access network swiftly with the available commodity devices may take much less time and effort than restoring the destroyed ones. Integrating both approaches allows the whole network to provide non-stop services on an end-to-end basis. We discuss four key technologies for improving the resilience of communication networks, namely survivable network deployment, demand-driven network resource management, ad hoc networking, and delay/disruption-tolerant networking.

3.1.1 *Survivable Network Deployment*

Research on network survivability dates back to graph theory [13], which states that a graph is k-vertex connected if and only if there are at least k vertex-disjoint paths between any two vertices. A similar result holds for edge connectivity. For the communication network to survive emergency events, it must be deployed with consideration of its robustness. The robustness of a network can be viewed as the cost or difficulty of destroying the desired connectivity of the network. The connection between two or more nodes in a network can be disrupted by destroying some network elements (nodes or links). In general, two links that are physically close to each other are more likely to fail together than two links that are far apart. Network designers should make the network as robust as possible under the given budget for network resources.The heterogeneity of modern networks means that traditional k-node connectivity and k-link connectivity-based approaches [14] are inadequate for modeling the robustness of modern networks. To model network failures in the event of emergencies, the notion of k-link connectivity should be extended to deal with cases where there is a cost for destroying a particular subset of network elements. This can be modeled as the probability of a subset of network elements being destroyed by an emergency event. The book by Frank and Frisch [14] contains an excellent discussion of most of the early works on the application of network flow theory to the design and analysis of what was then called invulnerable networks. Subsequently, significant advances of both theoretical and practical values have been reported. These developments are the result of advances in techniques for algorithm design.

Additionally, since it is not realistic to deploy a network anew, a practical method is to augment an existing network by adding new network elements. For example, satellite links may be added to significantly improve the robustness of a network, since the satellite links are less likely to fail together with the links in the ground in the event of a disaster. However, augmenting a network incurs a cost. We will study efficient ways to maximize the robustness of a network under given resource constraint. Since our network-upgrading problem is an offline optimization problem, combinatorial optimization approached will be used. We will study generalizations of these approaches to the network- upgrading problem here. We need study efficient algorithms for computing the $s-t$ tolerance of a given pair of nodes s and t, given the information on the cost for destroying a subset of network elements. We can estimate this information using historic data and the prediction of future disasters. Then, we can apply network flow-based technique to compute the $s-t$ tolerance of a given pair of nodes.

Redundant network construction for easy protection and restoration has also been studied extensively, such as using redundant tree [15] or disjoint paths [16]. By constructing a pair of redundant trees, called red and blue trees, it is able to guarantee fast recovery from any single-link/node failure in the network, if the failed node is not the root node [17], [18]. There are also studies on approximation algorithms for augmenting a wireless sensor network to meet connectivity and survivability requirements [19], [20].

3.1.2 Demand-Driven Network Resource Management

Handling the extraordinary traffic demands that occur during emergency events is another important issue. The damage caused by an emergency results in a shortage of computing and communication resources. Maintaining connectivity and expanding the capacity of the network, for both the stricken area and other areas, are essential requirements for big data-enabled decision-making.

As an example of such limited resources in the event of an emergency, consider the power supply for BSs. Some BSs may have no power supply following an emergency, or they may be equipped with batteries or a green power supply. Thus, resource management of these BSs must take into account the power constraint. However, such constraints should be removed after the grid power supply is restored. Uchiyama et al. studied energy resource management at BSs using cell zooming [21].

If the radio access network can be divided into a centralized control component for processing base-band signals and a distributed component for radio elements (i.e., antennas or physical cell towers), the requirements for the radio elements will be much simpler than for current BSs. Such software-defined radio access network architectures would enable power saving at location-associated communication facilities, and provide promising solutions for sustaining communication capabilities over very large geographical areas during a disaster. Previous studies [22], [23] discussed the possibility of applying the software-defined network concept to wireless cellular networks. By separating the control plane and the data plane of radio access networks, it may be possible to solve the problems of bandwidth allocation, interference mitigation, handover, and load balancing among multiple cells. However, using software-defined radio access network technologies for managing situations during emergencies is still an open problem [24].

New architectures enabled by software-defined network technologies, besides their possible use in maintaining network connectivity, also offer a promising approach for hosting heavy traffic during emergency events. For example, the value per bit of the information to be processed or transferred (in general, the value per unit resource) will become higher for people in stricken areas. Methods for allocating computing and networking resources should be adaptively and optimally tuned in response to emergencies. Thus, implementing new architectures to allow computing and communication systems to work in a flexible and scalable way is indispensable for emergency preparedness.

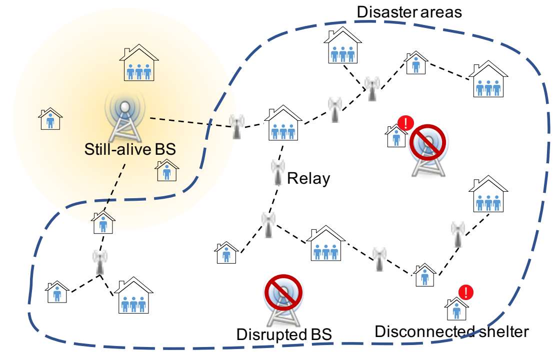


Fig. 3: An example of ad-hoc networking for providing connectivity in the disaster area.

3.1.3 Ad-Hoc Networking

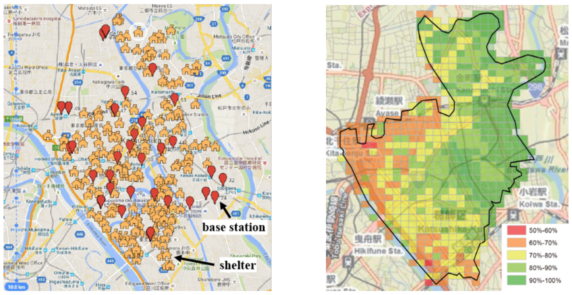
This technology deploys provisional access networks that provide users with communication environments they are accustomed to, even during emergency conditions. An access network is usually located within a small area close to the users' locations, with no redundant network resources. In this case, constructing a provisional access network swiftly with the available commodity devices and/or fast deployable emergency resources may take much less time than restoring the destroyed access network.

Smartphones, as a most popular and available commodity device, have frequently been adopted for provisional ad-hoc networks in disaster scenarios [25]. These networks can be quickly setup by ordinary users via on-site commodity mobile devices, without any requirement for additional equipment, by connecting them to the nearest Wi-Fi access point to form a multi-hop network for Internet access. However, this approach is highly limited by the distribution and mobility pattern of mobile devices (i.e., users) after a disaster, as most users will have gathered at specific locations, such as shelters, to obtain the necessary supplies [26]. Based on the multi-hop networking of users' devices, it is difficult to interconnect the shelters to the Internet [27].

The approach of using dedicated devices, such as deployable wireless relays, to recover network coverage in important area such as shelters has received considerable attention [28]. A field study was conducted with a vehicle-mounted base station and many dedicated low-cost wireless relays [28]. The results show that the network can provide satisfactory coverage to a large area via wireless relays and mobile devices. In such networks for disaster recovery, a natural challenge is the optimal deployment of a limited number of relays to extend the network coverage to disconnected locations, such as shelters, where most of the people have gathered after a natural disaster.

3.1.4 Delay/Disruption-Tolerant Networking

To overcome the challenge of intermittent connectivity between nodes in ad-hoc networks, the paradigm of Delay/Disruption-Tolerant Networking (DTN) was proposed. In DTN, mobile nodes carry the packets (or data bundles) until they encounter other nodes, and choose whether to forward the packet at these encounters. This store-carry-forward approach tries to mitigate the problem



(a) Shelters and BSs (b) Network availability

Fig. 4: A real-world scenario for performance evaluation.

of partitioned networks. Most of the proposed DTN routing protocols focus on enhancing the delivery ratio and/or reducing the overhead [29]. This is largely achieved by using past encounters as a metric for calculating future delivery probabilities. In an emergency scenario, DTN nodes such as rescue vehicles and unmanned aerial vehicles (UAVs) are usually administrated by the same authority. This suggest another research topic regarding the mission/trajectory planning that most improves the DTN routing performance under some guaranteed delay and delivery ratio [30].

3.2 Real-world Application and Case Study

Real-time and reliable data collection from emergency sites is the key for data-driven emergency management. To achieve this, maintaining network connections for those affected by the emergency event is vital. In the communication network infrastructure, the access network elements are usually the most vulnerable part following emergency events such as natural disasters, as these network elements (e.g., BSs) are geographically distributed in hostile environments and are prone to damage. Unlike the backbone network, it is extremely hard to deploy access networks with sufficient redundancy for better survivability. Therefore, a quickly deployable ad-hoc network is the primary solution to this problem.

We demonstrate a relay deployment technique for a wireless ad-hoc network that ensures the network connection is efficiently restored to people in the affected area [31]. Practically, there are many geographically distributed shelters, such as schools and community facilities, where people tend to gather. The capacity of each shelter is different, so the accommodated population is different. An example of disaster recovery ad-hoc networking is illustrated in Fig. 3. The relays could be small, low-cost, self-powered devices, which can be deployed almost anywhere. For instance, a fast deployable Wi-Fi relay module with a solar power panel the size of a suitcase has been developed [28].

To facilitate information collection and dissemination in a disaster scenario, the first responders are frequently faced with the problem of recovering the network connection as widely as possible under limitations such as the number of relays, communication distance, and population distribution. This can be mathematically formulated as a graph problem, and the corresponding algorithm is proposed. The effectiveness of the proposed solution was evaluated in the real-world scenario illustrated in Fig. 4. The availability of the BSs is adopted from a study of data-driven network availability [32]. Fig. 4b shows a snapshot of network availability in Katsushika ward following an earthquake in the Tokyo area. The population distribution is shown in Fig. 4a, and this can be further accurately estimated using the method described in [33]. Through the case study of a wireless multi-hop networking deployment, we are able to collect essential data and disseminate information to people in the emergency area, which is necessary for efficient data-driven emergency management.

4 Mobile Communication Networks for Emergency Management

Large-scale disasters, such as devastating earthquakes, floods, wildfires, and tsunamis, can result in massive blackouts and cause severe damage to telecommunication infrastructure. It usually takes months to repair the damaged infrastructure. Therefore, ECNs [3] are needed to reduce loss of lives, limit damage to people and property, and satisfy the sharply growing communication demands of disaster victims. ECNs must be designed to offer reliable post-disaster communication [34]. For example, an ECN center can collect messages from disaster areas and inform the victims about the overall situation and appropriate actions.The study on resilient communication networks aims to safeguard the communication of victims in a wide regional level. However, to the individual of victims, especial to whom living in remote areas, a refined supplement of the ECN will be the key technology to improve the efficiency and fairness in relief. Therefore, besides the macrolevel resilient communication networks, microlevel and crowdsourcing communication networks for emergency management are also necessary for reliable relief.

Nowadays, smartphones are ubiquitous in our daily lives, and each of them has embedded sensors such as GPS, cameras, compass, gyroscopes, microphones, and light sensors. Hence, when a group of smartphones connect to the Internet through cellular or Wi-Fi networks, or when they connect to one another in a cluster using Bluetooth, a sensor network emerges that can be exploited to organize ECNs for disaster-relief tasks. Smartphone-based ECNs have therefore attracted considerable attention in recent years.

Fig. 5 shows a typical smartphone-based ECN. After a disaster, the affected area is often split into isolated communities. As the telecommunication infrastructure is frequently damaged, these communities may be disconnected from both the Internet and one another. In such cases, as illustrated in Fig. 5, techniques for connecting the networks include vehicular mobile stations [35] and DTNs [36]. In areas that remain accessible to vehicular mobile stations, people can use social media apps such as Twitter and Instagram to communicate with family and friends through the Internet connection provided by the mobile stations. In road-blocked areas, community residents could instead share content with their smartphones based on the UAVs and Device-to-Device (D2D) technologies. Smartphones can also collect data from IoT devices located in the disaster area. The collected data can be transmitted by mobile stations and aggregated in remote cloud databases for big data analytics. The useful

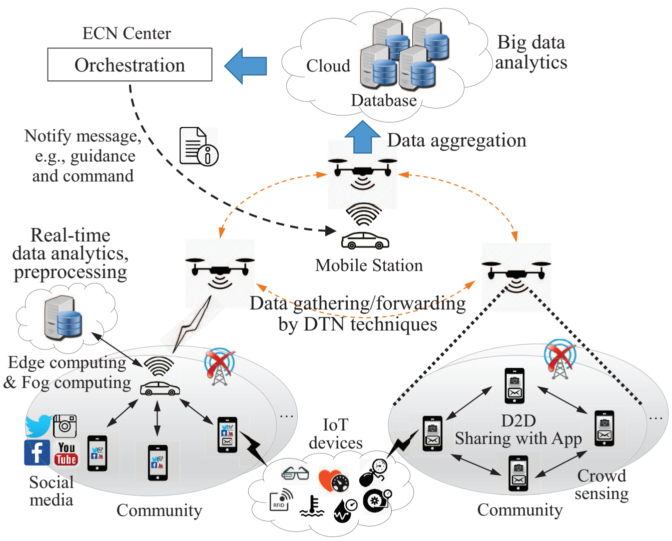


Fig. 5: Overview of a smartphone-based ECN architecture.

information extracted in this way is delivered to the ECN center, where it is used to make disaster-relief decisions and send notifications back to the affected areas, e.g., evacuation guidance and rescue commands. In edge networks, where edge servers have been deployed, the real-time big data preprocessing and analytics can be accomplished through edge computing [37].

In recent years, there are several key concepts for ECN appearing in the emergency management field:

Situation Awareness: Situation awareness information plays an important role during disaster-relief operations, because the ECN center makes rescue plans based on road damage, population distribution, resource requirements, and medical demands. Thus, timely situation-awareness information is critical for ECN management. Leveraging the pervasiveness of sensor-equipped smartphones, *opportunistic sensing* technology [54] is a promising paradigm for scalable context monitoring, e.g., for sensing the behavior of large crowds or monitoring the environment. In particular, *crowd-mobility* studies, such as the prediction of crowd mobility in public areas [36], are highly useful for guiding evacuation and preventing casualties caused by chaos and panic in the crowd.

Disruption/Delay Tolerant Networks: As mentioned already, the communities in a disaster area may be split into multiple isolated ECNs. DTNs attempt to provide end-to-end connections across communities by exploiting mobile stations and aerial vehicles [35].

Big Data Analytics: A large volume of data can be collected from various sources, such as IoT devices located in disaster areas, smartphone-based networks, and social-media networks. To leverage the collected data and better understand the situation, *big data analytics* is essential. Recent studies [36] have analyzed big datasets from social-media sources such as Twitter and Facebook to better respond to disasters and manage emergency networks.

Edge and Fog Computing: By 2020, an estimated 50 billion devices will be connected to the Internet [55]. This IoT will generate a tremendous volume of data that needs to be processed and analyzed. Fig. 5 shows how the IoT devices located in a post-disaster area and the smartphones themselves yield a large volume of raw data. To retrieve useful information for rescue teams in real time, distributed on-site processing is required. However, the conventional cloud-based data processing paradigm directs the data streams to remote cloud servers for processing, which may result in network congestion and traffic delays. *Edge computing* [37] and *fog computing* [56] reduce the data transmission requirements and processing load on the cloud servers by processing data in computing nodes located at the edges of networks. This edge computing near the IoT devices and fog computing near the users' devices will most likely be located in the same area in a disaster situation. Thus, edge computing and fog computing can reduce situation-awareness delays and improve the quality of service in smartphone-based ECNs.

4.1 Methodologies and Key Techniques

We proceed to review state-of-the-art existing methods and key techniques for smartphone-based ECNs.

4.1.1 ECNs based on Ad-hoc Networks and DTNs

To ensure a rapid response to disasters, recent studies [36], [38], [39], [40], [41] explored the use of ad-hoc networks, opportunistic networks, and DTNs to construct ECNs in disaster areas for the rescue, relief, and evacuation of victims. For example, Trono et al. [41] used DTN communication to develop a smartphone application called DTN MapEx, which generates and shares maps of disaster areas by exploiting multiple nodes in the system. This application minimizes the computational workload of individual devices, because map generation tasks are shared between the mobile sensing nodes in the DTN. Higashino et al. [36] investigated disaster mitigation, leveraging DTN-enabled distributed micro-modules to design a smartphone-based crowd-event detection architecture.

4.1.2 Mobile Base Stations

To foster situational awareness about disasters, ECNs must support data capturing and aggregation. A handful of studies [35], [42], [43], [44], [45] apply vehicle- or aerial-based mobile BSs for these purposes. For example, Gomez et al. [42] considered the ABSOLUTE project [46], and developed a prototype low-latency IP mobile network with wide coverage by combining aerial, terrestrial, and satellite communication networks. The aerial BSs acted as the central components for providing a resilient communication service to mobile devices. A low-cost balloon-based network [44] has also been proposed for post-earthquake rescue. Li et al. [43], [45] developed a disaster management network based on mobile stations implemented by drones and vehicles equipped with sensors and network interfaces. The aim was to support disaster management tasks by sensing damage conditions, collecting information, and delivering messages to disaster areas. Narang et al. [35] proposed a cyber-physical buses-and-drones-based mobile-edge infrastructure for emergency communications following large-scale disasters in which the cellular infrastructure has been destroyed.

TABLE 1: Key Techniques and Methodologies for Smartphone-based ECNs

Contribution	References	Key Techniques and Methodologies
Construct ECNs	[36], [38], [39], [40], [41]	ECNs based on ad-hoc networks, opportunistic networks and DTNs
Data collection and gathering	[35], [42], [43], [44], [45]	mobile base-station based mechanisms
	[46], [47], [48]	Device-to-Device (D2D) communication based mechanisms
	[49], [50], [51], [52], [53]	crowd-sensing based mechanisms

4.1.3 D2D Communications

Many recent contributions have explored how D2D communications can extend network coverage in the context of disasters. For example, in the ABSOLUTE project [46], short-distance D2D communications are provided for rescue teams and emergency agencies when the conventional network infrastructure has been damaged by a disaster. Orsino et al. [47] studied social-aware data collection and information diffusion using D2D communication techniques, proposing an approach that can be applied to emergency networks for public safety. Based on D2D communications in an ad-hoc network, Meurisch et al. [48] proposed an emergency communication system called NICER911, which aims to provide reliable communication and emergency services in disaster areas with compromised infrastructures.

4.1.4 Crowd Sensing

To enhance situation awareness during disasters, several studies [49], [50], [51], [52], [53] have developed elaborate smartphone-based crowd-sensing techniques. For example, Higuchi et al. [49] proposed a low-power collaborative localization algorithm that captures the stop-and-go behavior of indoor pedestrians. Based on the cooperative operations among multiple smartphones, Noh et al. [50] developed an infrastructure-free localization technology with high-speed positioning functionality. Kojima et al. [51] proposed a new application that estimates the reason for particular human crowd events using mobile crowd-sensing techniques. To improve bandwidth utilization and reduce energy consumption, Zuo et al. [52] explored an image sharing mechanism that promotes on-site situation-awareness about disasters such as earthquakes and typhoons. The shared images are collected via smartphone-based crowd-sensing techniques. Because cameras can help rescue teams gain situation awareness, e.g., about trapped victims, Dao et al. [53] implemented a network of smartphones with cameras to transmit pictures in an energy-efficient manner.

4.2 Real-world Application and Case Study

We now review some real-world applications of smartphone-based ECNs, before presenting a disaster-management case study.

4.2.1 Dedicated Smartphone Apps For Disaster-Relief

Peng et al. [57] developed a Bluetooth-based smartphone app called E-Explorer, which can deliver rescue information to survivors trapped in post-earthquake sites. To better support emergency communication and the rapid investigation of earthquake damage, Han et al. [58] extended the iOS-based E-Explorer application to other platforms. For more efficient responses to disasters such as large-scale earthquakes and tsunamis, Miyazaki et al. [59] developed a resilient information management system that runs on both Android and iOS and provides convenient information management and data exchange functionalities to rescue teams and victims following disasters.

4.2.2 Case Study: RIM System

We now describe our ongoing project, the Resilient Information Management (RIM) system [59].

The RIM system uses smartphones and UAVs to provide delay-efficient and reliable solutions in harsh disaster environments where conventional cellular communication infrastructures have become unavailable or severely damaged. Although an ad-hoc mobile social network built through mobile devices such as smartphones may be the most straightforward communication approach, the resulting delivery delay might be too long for time-critical disaster-relief tasks.

As shown in Fig. 6, we propose to reduce delivery delay by using UAVs with wireless communication capabilities. The UAVs follow designated routes to collect information about damage, injuries, and medical demands from specified sites, and deliver the gathered data to an information center. In this way, delivery latency can be greatly reduced compared to systems that rely on ad-hoc communication.

In this mobile devices-based architecture, we have designed an integrated information management and sharing mechanism that benefits both rescue teams and victims [59]. An online algorithm that dynamically schedules mobile stations based on management task priorities has also been proposed [43]. Further work on the RIM system will study, for example, energy-efficient scheduling of drone routes during the data collection missions.

5 Remote Sensing Technologies for Emergency Management

Remote sensing is the science of gathering sensor information about objects or regions from a distance, usually from satellites. Remote sensing applications are often deluged with enormous volumes of remote sensing data, and can thus be considered as typical data-intensive systems. Remote sensing data, namely satellite remote sensing big data, have several peculiar characteristics in terms of their diversity and high-dimensionality. During natural disasters, large-scale climate monitoring applications can process local-to-universal multi-temporal and multi-sensor remote

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2020.2972871, IEEE Transactions on Big Data

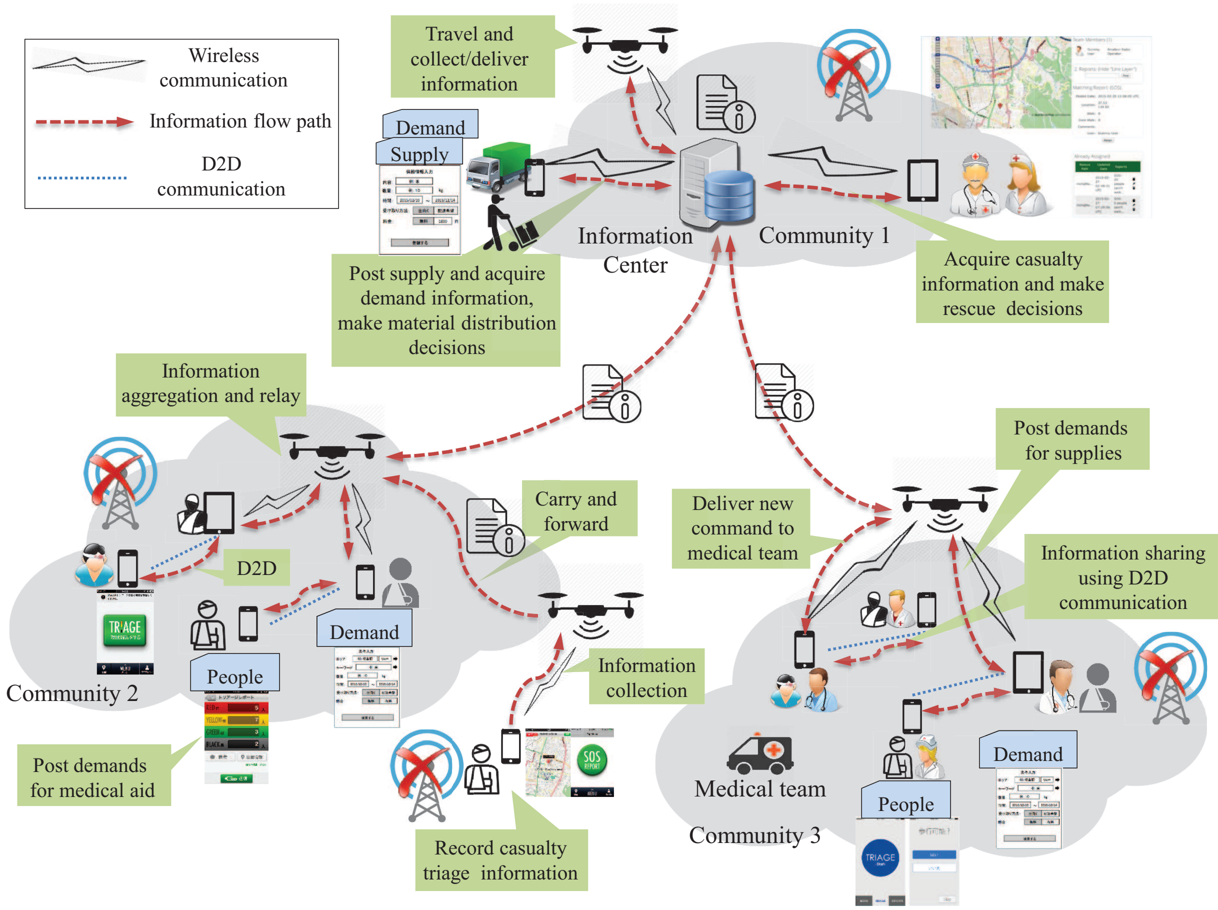


Fig. 6: Case study: the smartphone- and drone-based Resilient Information Management (RIM) system.

sensing data. In this way, remote sensing systems have demonstrated their potential utility in delivering critical information for emergency management.

5.1 Key Technologies and Methodologies

Large-scale satellite imagery data mining: One of the origins of big data for alert forecasting is from satellite imagery. Standard satellite image data can be used for large-scale emergency impact assessment by mining the spatial scale of influence of a disaster. The satellite types and the spatial and temporal resolutions required to assess emergency damage are reviewed in [60]. For satellite data to be valuable in the aftermath of an emergency, processes to gather imagery of the emergency incident, analyze the received information, and equip technical staff to carry out the task should be established well in advance [61]. To achieve this, the efficient processing technology for satellite imagery data is necessary. In recent years, the sparse representation gives interesting results for discarding remote-sensing dynamic datasets [62].To sparsely represent the spatiotemporal remote-sensing Big Data, a new dictionary learning algorithm is proposed here by extending the classical K-SVD method. The information, from both old and new samples, is explored in the proposed incremental K-SVD (IK-SVD) algorithm, but only the current atoms are adaptively updated. This makes the dictionary better represent all the samples without the influence of redundant information from old samples. Additonally, many nonlinear time series methods and dimension-reduction methods have been developed which could be applied to big data with high-dimensional characteristics and improve the computational efficiency. [63].

Camera data mining: Mining the video and imagery data which are from surveillance cameras will be the most direct method to detect the localised people evacuation and the disaster situation. Currently, the camera data-based risk assessment method is an emerging trend in the emergency management field which is able to overcome the tradeoff between accuracy and efficiency. Deep learning-based on ex-ante online risk assessment was proposed by [64] and implemented in a library. Model sensitivity analyses and large-scale tests demonstrate the usability and superiority of the proposed method. In addition, although several early-warning solutions rely on automated modeling and algorithmic approaches, crowdsourcing can augment attempts to sieve the signal from the noise in large-scale data. Volunteers have been requested to geotag and classify pictures of affected people and objects in post-emergency analysis [65]. TomNod and the OpenStreetMap have employed digital technology to help emergency authorities and volunteers [66].

Multi-model integration: Big data analytics can assist in the mitigation phase of risk and vulnerability analysis. Lucid visualization techniques such as "Hotspot" maps and hexagonal cells can blend the value of biophysical and social data to prioritize affected areas. Climate model results and emergency risk models can be integrated with remote sensing images to determine societal and monetary exposure to risk [67]. The main benefit of satellite data is that it is gathered over time, permitting for the mechanized evaluation and updating of risk models at regular periods. For example, the Global Inundation Extent from Multi-Satellites (GIEMS) [68] provides a low-resolution scheduled dataset of surface water extent and dynamics. Recently, big datasets have been tested to increase the size of potential variables in risk mapping, along with climate models and satellite pictures. For instance, spatial video has been used in Haiti to rapidly identify locations inundated by sewage and water for cholera hazard mapping [69] and in Los Angeles to determine exposure to wildfire [70]. Nowadays, data from drones and "high-res" satellite images can be combined with crowdsourcing to yield comprehensive structure maps of buildings (Opencitiesproject.org).

5.2 Real-world Application and Case Study

Several well-known mapping apps use remote-sensing big data, such as Google Maps and Google Earth. These apps utilize the Google Earth Engine, which brings together a multi-petabyte repository of satellite imagery and geospatial datasets with planetary-scale analysis features to detect changes, map trends, and quantify dissimilarities on the Earth's surface.

Urban areas facilitated with sensors have an advantage during emergencies, because actionable information can be gathered to help decision-makers. Moreover, machine learning techniques can be used in the analysis of the sensing images. Machine learning algorithms enable the generalization and prediction of observations. Crowdsourcing is an innovative perspective for emergency management. It uses satellite data, citizens sensors, and volunteer mappers from the affected area to furnish real-time data and analysis during emergencies.

There are many emerging applications in this field, such as a new algorithm that uses time-series analysis to create images based on LandSat data [71]. This algorithm has achieved stimulating results in bridging gaps in knowledge. Moreover, deep learning, which learns the characteristics and discriminative features of data in a hierarchical way, has also been used to analyze remote sensing data [72]. A Hessian-based method [73] has also been applied to uncertainty quantification in emergency situations.

An example of the crowdsourcing of maps, social media data, human sensing data, and satellite data for emergency management occurred after the flooding in western Norway in 2014. Western Norway is frequently referred to as fjordland after the region's most striking natural feature. Winter lasts from mid-November until April, and climate situations during this season fluctuate extensively based on the geography of the area. There was heavy rainfall in western regions of Norway during autumn 2014, leading to hazardous flooding in several places that washed out roads, bridges, and countryside.

When rivers overflowed in the counties of Sognog Fjordane and Hordaland on October 28, 2014, sensing images and maps of the region were collected. Authorities had to locate those affected by the flooding to provide aid and emergency response. The ability to crowdsource the data gathering and mapping processes was important considering the urgency of relief operations. Volunteers and local authorities gathered satellite images for tracing and recording the streets, buildings, and other objects of interest.

Big data analytics processed massive volumes of social media data, such as tweets, and clustered them according to the issue (high content overlap), zone (for posts with GPS tags), and time span. Clusters of tweets are the result of high social network activity in a particular area. Furthermore, machine learning classifiers were used to automatically identify clusters that were likely to be of interest to emergency service providers.

Moreover, crucial details about river flooding conditions and damage control activities were used to schedule flood control actions. To offer the emergency service providers precise and multi-temporal data, high-resolution sensing data were utilized to plot the river flooding conditions and ongoing hazard control activities. For some flood-vulnerable regions, post-emergency sensing data were analyzed to record the alignment of the river along with hazard control defenses.

6 Human Mobility and Urban Sensing for Emergency Management

In the course of emergency events, it is vital to understand, model, and predict the population behavior and mobility at a citywide level so as to plan effective transportation scheduling, humanitarian relief, and post-disaster management. For instance, the Great East Japan Earthquake in 2011 disrupted the public transportation systems in the Greater Tokyo Area, the largest metropolitan area in the world, and caused widespread traffic chaos and urban disorder. If city managers could accurately predict how a large number of people move and select their transportation mode, they would be able to plan effective transportation scheduling and avoid overcrowding or traffic jams. A dazzling light show in Shanghai on new year's eve, 2014, attracted a huge crowd to celebrate the arrival of 2015. Unfortunately, the crowd density was soon out of control and a tragedy ensued: 36 people died and 47 were injured in a stampede. The Shanghai local police agency admitted that the crowd density was underestimated. Reconsidering this event, it would be useful to provide an accurate forecast of the crowd density before it became too dangerous. Obviously, if we could observe and predict human mobility and density in advance, suitable countermeasures could be enacted to prevent similar tragedies from happening.

Understanding and modeling human mobility at a citywide level is a core research direction for emergency management and urban computing. However, this research is difficult because there is no reliable method for sensing human mobility. Recently, however, with the rapid development of mobile internet technologies, massive location acquisition and human mobile sensing data are being continuously generated from various sources, such as smartphones, GPS devices on cars, WLANs, IC cards, and location-based social networks. This big data offers a new way to circumvent the problems of previous research for understanding human behavior, because it is instantaneously available, offers high spatial and temporal resolutions, has no interview bias, and provides comprehensive data for large populations [74], [75], [76], [77], [78], [79]. Against this background, accurate predictions of citywide human mobility become possible, and these are critical to many intelligent urban systems, such as traffic regulation and crowd safety surveillance, used for emergency response and disaster management.

However, influenced by a variety of factors (e.g., weather, urban structure, event type, traffic conditions), citywide human mobility modeling remains highly complex in the real world. The main complicating factors are as follows: (1) A modern city usually contains huge road segments and a complex transportation network, representing a highly nonlinear and complex system. It is very difficult to exploit heterogeneous human sensing data and properly generalize the target conditions. (2) Human behavior and mobility patterns usually have a high degree of freedom and variation, and are impacted by many factors. Thus, human mobility and behavior are very difficult to model and predict. (3) The urban areas of modern cities have a very large spatial domain, making accurate predictions difficult. For instance, the Greater Tokyo Area extends over 3.925 km^2, and the larger metropolitan covers 14.034 km^2. (4) To build an effective human mobility prediction model or mine hidden human behavior patterns, we usually have to collect sensing data over long time periods. The large temporal scale usually makes the entire modeling process very complex. (5) Most emergency or rare events only influence human mobility in specific regions. For example, a traffic accident may only change the local human mobility patterns and the key paths passing through the event. Therefore, if we labeled a rare event at a citywide level, many irrelevant subjects may become involved (e.g., someone in a nearby shopping mall may be labeled as part of the "traffic accident"). Thus, attractive though it is, accurate predictions of human mobility remain a significantly challenging research topic.

6.1 Key Technologies and Methodologies

In recent years, a vast number of studies have analyzed human mobility big data, such as GPS logs [80], [81], [82], social network data [83], [84], query data from routing apps [85], IC transport card data [86], and bike rental data [87]. These massive datasets have great potential to solve urban problems [6], especially for emergency response and disaster management [88], [89], [90]. Lu et al. [76] collected mobile phone data from 1.9 million people in Haiti and analyzed the population displacement during the 2010 Haitian earthquake. Song et al. [91] collected data from 1.6 million GPS users in Japan to understand and predict human mobility and evacuation behavior during the Great East Japan Earthquake and Fukushima nuclear accident in 2011. They demonstrated that human mobility after large-scale disasters was more predictable than previously thought.

Individual Human Mobility Modeling: Most of the above models struggle to predict the mobility or behavior of individual people accurately. Thus, Song et al. [88] proposed a hidden Markov model-based predictor to forecast human mobility during a natural disaster. They extracted important locations such as "home," "work," "social," and "unknown" for fine-grained citywide human mobility prediction. Auxiliary data indicating the key factors influencing human mobility are also required, which makes this unsuitable as a general human mobility predictor that can address both routine human mobility and behavior in the event of an emergency. Yabe et al. [92] extracted 18 features and applied logistic regression to predict the occurrence of irregular human movement during frequent mid-level disasters. Fan et al. [80] used a finer location representation by discretizing the Tokyo region into meshes and representing locations as a mesh code. Only the most recent trajectories were utilized to make future predictions. This approach does not require any labeling work and can be easily implemented as a general human mobility predictor, but it discards all historical information, which leads to a loss of accuracy. Another direction is to use auxiliary data to make long-term predictions and provide early warnings about crowd safety. Konishi et al. [85] leveraged transit routing queries to identify long-term human mobility related to specific events, and [93] explored the probability of using map query data to prevent accidents such as the Shanghai stampede.

Human Mobility Flow Modeling: Understanding the basic patterns of a very large number of population movements is also important for urban emergency management. Understanding the flow of people [94] and recommending location-based services [95], [96] use tensor factorization to decompose urban human mobility. Traffic flow can be seen as a special human mobility constraint on road networks, and this has been studied for traffic flow prediction [97] and traffic congestion monitoring [98], although some kind of initial individual model is required for each road segment. Moreover, Song et al. [74] explored the upper bound of the predictability of human mobility, and Zheng et al. [99] proposed an unsupervised learning algorithm for location prediction. A more advanced trajectory calibrating algorithm was proposed in [100].

Deep Learning Technology for Human Mobility Modeling: More recently, deep learning technology has shown great potential in the fields of natural language processing and computer vision. There have been some attempts to introduce deep learning into human mobility prediction. Song et al. [101] proposed a multi-task long short-term memory deep learning framework to predict both the transportation mode and location. These two tasks are highly correlated, and can thus boost each other's performance. Residue networks and convolutional neural networks have been successfully used to predict citywide crowd flows [102], and a long time-series model for predicting the density value of each grid cell has been constructed [103]; in contrast, our approach aims to predict citywide human mobility through sequence classification. In addition, researchers have applied deep learning to traffic problems in which the traffic speed and transportation mode are associated with human mobility [104], [105], [106], [107].

6.2 Real-world Application and Case Study

Researchers at The University of Tokyo developed the DeepMob disaster management system [90] for effectively

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2020.2972871, IEEE Transactions on Big Data

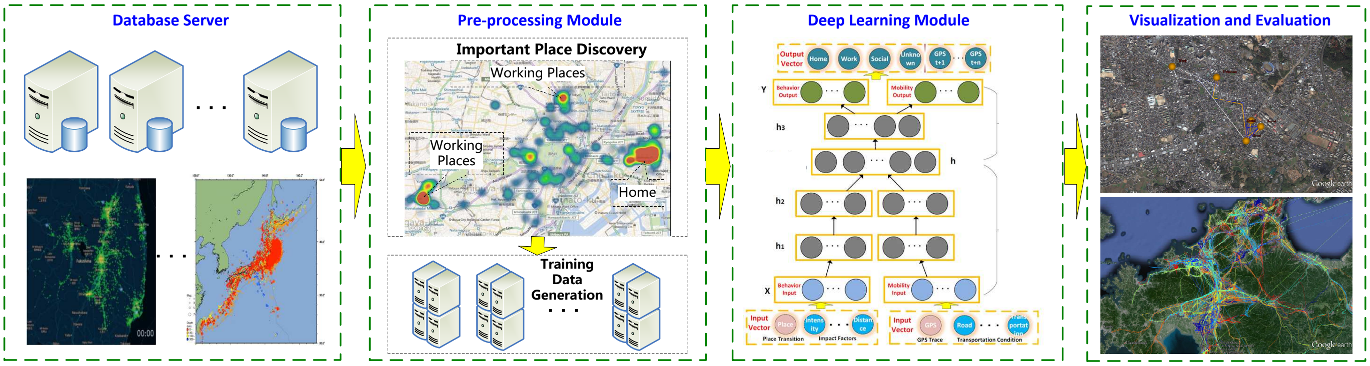


Fig. 7: DeepMob: DeepMob [90] is disaster management system for effectively predicting and simulating population evacuation behavior and mobility following various types of earthquakes.

predicting and simulating population evacuation behavior and mobility following various types of earthquakes. Users input observation data (e.g., GPS traces), disaster information (e.g., earthquake magnitude, earthquake intensity scale), and transportation conditions, and the system automatically predicts and simulates the evacuation types and travel routes in subsequent time steps.

DeepMob relies on heterogeneous big data sources, e.g., GPS records, transportation and road network data, Japan earthquake data. To discover knowledge from these data, the learning architecture performs two tasks: (1) learn the deep feature representation from disaster information data and people's location transitions (as shown in Fig. 7-a); (2) learn the deep feature representation from GPS traces and transportation network data (as shown in Fig. 7-b). To enhance the system performance, DeepMob performs multimodal learning [104], [108], [109], [110] to learn both the feature representations and deep knowledge about population behavior and mobility following different types of earthquakes (as shown in Fig. 7-c).

Experiments show that DeepMob can achieve 87.8% accuracy in predicting people's evacuation types and 79.57% accuracy in predicting their evacuation routes following various types of earthquakes.

7 Online Social Network Analysis for Emergency Management

Social Media. The emergence of the Web 2.0 paradigm has led to the widespread adoption of technology platforms for content generation and sharing. These platform applications enable computer-mediated communication among citizens to create and share information online, and to pursue topical interests by joining online communities [112] and networking with like-minded users. The growth of social media platforms to fit the interests of different users combined with the worldwide adoption of mobile technology has made the role of social media ubiquitous in our daily lives. For instance, while only 5% of Americans used some form of social media platform in 2005, some 69% of the population were using social media in 2016 [113]. Social media has created an opportunity for citizens to act as citizen sensors [114], which can be extremely valuable for the sensing and sharing of useful observations during emergencies. A citizen-driven information infrastructure has created a new information-sourcing channel for emergency management organizations, providing enriched information for dynamic situational awareness and improved response services [115] [7]. During recent emergencies and disasters, social media messages have included relevant information such as caution-advice and damage reports [116], requests for and offers to help [117], and emotional support for the affected community [118]. However, the relevant information is often buried within a vast quantity of noisy, large-scale unstructured data. A variety of multimodal (text, images, videos) data are generated at high velocity, presenting emergency management organizations with a big data problem [119].

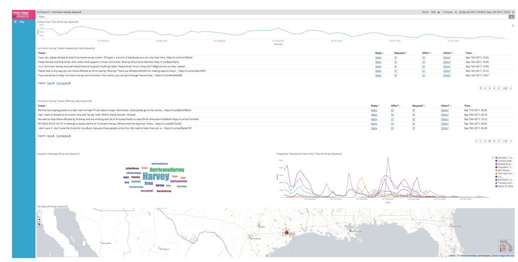


Fig. 8: Example of a customized dashboard *CitizenHelper* [111] during Hurricane Harvey to visually interact with relevant social media messages by selecting time (widget 1) and location (widget 4), any trending hashtags (widget 5) or trending mentioned users (widget 6) as well as filtered messages of public or groups requesting and offering help (widgets 2 and 3), extracted by active learning techniques in the background that have an ability to take feedback (e.g., irrelevant request) from a viewer using buttons next to the tweet message.

Crowdsourcing on the Web. Crowdsourcing [120], [121] has become a very popular concept in recent years. However, its roots come from different concepts in diverse areas of study, such as peer production, collaborative systems, and collective intelligence. Crowdsourcing systems combined with the power of digital connectivity over the Web have revolutionized distributed information processes. While there is no specific definition for the concept, according to Doan et al. [122], "[a] crowdsourcing system enlists a crowd of users to explicitly collaborate to build a long-lasting artefact that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2020.2972871, IEEE Transactions on Big Data

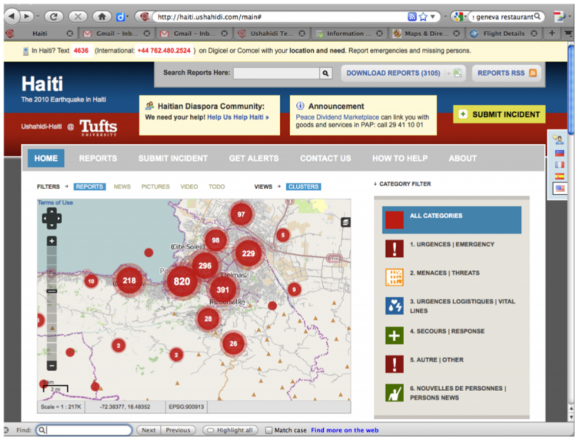


Fig. 9: Illustration of Crisis Mapping during Haiti that was supported by crowdsourced reports [125].

is beneficial to the whole community." While social media has demonstrated potential for collecting partial large-scale information during emergency events, the crowdsourcing phenomenon has demonstrated the power of the masses to help process this big data into coherent products, e.g., annotating data for developing machine learning classifiers [123]. Furthermore, crowdsourcing has initiated various volunteer and technical communities, such as members of the Digital Humanitarian Network [124], which activates and performs distributed tasks for accomplishing the overall goal of serving an information need of an emergency management organization.

7.1 Key Techniques and Methodologies

There are several applications of big social data for all phases of the emergency management cycle. In the past decade, with the rising adoption of social media, a variety of computational techniques in areas such as data mining, machine learning, natural language processing, and network sciences have been enabled by the easier access to data, such as public Twitter streams. We now describe some of the key problems and the techniques that can be used to address them (for comprehensive surveys, see [116], [119], [126]). The basic requirement for any social media mining task applied to emergency management is the early identification of the incidents; therefore, event detection is one of the critical techniques [127], [128]. For instance, Sakaki et al. [127] showed the efficacy of Twitter in detecting earthquakes in real time. After event detection, one needs a relevance criterion for information filtering, such as a relevant keyword set, so that only event-related social media messages are collected from the large amount of noisy, operationally irrelevant, content shared on social media. Given that manual keyword sets can be biased and become outdated during the rapidly changing events following a disaster, domain modeling and topic tracking are important techniques. Therefore, there have been attempts to adapt existing domain models or dynamically create models for event relevance that help identify and filter relevant social data for analysis [129], [130]. Once the data have been collected, either in streaming mode or batch mode, their processing requires a variety of techniques to extract information that will improve situational awareness and decision support. These techniques are primarily content-based, network-based, user-based, context-based, or employ visual analytics for the ultimate human–computer interaction.

Content-based techniques explore different information types in the message content [131], e.g., classifying and extracting topics of interest such as caution-advice and damage reports [116], modeling behavior such as requests for and offers to help [117], [132], [133], and measuring credibility and detecting rumors [134], [135], [136].

User-based techniques focus mainly on the identification of a variety of user categories [137], such as on-the-ground informants [138], emerging informants [139], influential users [140], real and virtual volunteers [141], and organizational users [142], [143].

Network-based techniques primarily investigate information diffusion for message reachability [144], [145] and community formation and evolution [146], [147]. In addition, simulation and agent-based modeling are useful methods for studying social network behavior before, during, and after disaster events [148], [149].

Context-based techniques help enrich the metadata of streaming data instances, such as geo-locations of the information source, which are typically present in less than 2% of the records. Geo-tagging [150], [151] and spatiotemporal analytics [152], [153] are examples of this type of technique for enhancing the modeling and analysis of social data.

Visual analytics techniques provide a means of easily understanding the information extracted from social data. The ultimate aim is to assist the emergency management teams and inform the public. One popular approach for social analytics visualization is a customized dashboard, as shown in Fig. 8 [111], [154], [155], [156]. This provides a spatiotemporal organization of information about what is happening where and what assistance is required. A key component of the dashboard is geo-tagged data visualization. This is often used as a standalone 'crisis map' [157], which has been an effective tool in various disasters over the last decade. Crisis maps provide geo-tagged information at one place and enhance situational awareness. Fundamentally, maps are a key provider of situational awareness in the incident management community, and new forms of information channels allow the crisis map inputs to incorporate social media and crowdsourcing feeds. Crisis maps are important information products that assist emergency responders in unknown regions, providing the right directions at the right time.

7.2 Real-world Application and Case Study

In a study [158], over 8.5 million tweets during Hurricane Harvey were analyzed to help to make decisions on disaster relief distribution. The geographic context was extracted from the tweets message content. Then, based on two data sets retrieved from authoritative data sources (The first is a digital elevation model obtained from the US Geological Survey for the derivation of the flow network. The second is the Land Use Dataset Council, available from the Houston-Galveston area, a consortium of Texas local governments, including Harris County, the key study area.), the comprehensive disaster relief location was detected. Finally, by employing a mixed-integer linear programming model and particle swarm optimization, the optimal layout of temporal rescue centers are decided.

Moreover, during the Haiti earthquake response, Ushahidi's Haiti Map [125] was commended by relief agencies for being extremely helpful in informing them of the situation on the ground (Fig. 9). Within the context of crowdsourcing efforts for emergency management, crisis mapping has recently emerged as a major open source technological concept for visually representing the crowdsourced information processed by distributed workers, creating a live crisis map for situational awareness. Starting with Ushahidi's project, which leveraged live collaborative mapping based on crowdsourced reports during post-election violence in Kenya in 2008 [157], crisis maps have become a core element of any emergency response that is assisted by citizen-led volunteer communities [124], [159]. Furthermore, volunteer and technical communities such as the Humanitarian OpenStreetMap team have become well-respected and reliable sources, rapidly providing a crowdsourced, collaborative geographical map during crises across the world to facilitate social data and crowdsourcing-driven maps [160].

8 Knowledge Graphs for Emergency Management

From the preceding sections, it is clear that emergency management relies on a broad variety of big-data sources including automated and human sensor data, mobility and communication data, and open and government data. In emergency situations, these and other sources must rapidly be made interoperable and available for complex querying, processing, and reasoning in ways that are not always foreseeable in advance [119]. *Knowledge graphs* originate from Tim Berners-Lee's vision of a machine-processable web of data that augments the original human-oriented web of documents [161], [162]. The central idea is to represent data semantically as graphs, with nodes that represent concrete objects, information, or concepts, and edges that represent meaningful relationships. The Resource Description Framework (RDF)[1] is the central standard for storing and exchanging knowledge graphs as files. They can be queried and manipulated using the associated SPARQL[2] language. The RDF Schema (RDFS)[3] and Web Ontology Language (OWL)[4] standards support even more precise semantics along with automated reasoning, in particular when the description logic subset (OWL-DL) of OWL is used.

These and other *semantic technologies* make knowledge graphs possible. Together, they offer principles and standards that can be used to fulfill central information interoperability needs in emergency situations. For example, [163] reports the development of SokNOS, an OWL-based disaster management system that offers better support for system extensibility, simplified database integration, improved search operations, external sensor discovery, plausibility checking, and visualization. In [164], the e-Response system, which integrates semantics-based and other tools to assist high-level tactical command during large-scale emergencies, is presented. e-Response uses knowledge graphs to encourage and facilitate collaboration between emergency operatives who may use different tools. Its use was illustrated through the example of a fire emergency scenario.

The Linked Open Data (LOD) principles [165] offer further advice for making data available as knowledge graphs. The central principles of LOD are as follows: (1) sharing graphs using standards such as RDF, RDFS, OWL, and SPARQL; (2) using Internationalized Resource Identifiers (IRIs) as standard names for nodes and edges; (3) defining these IRIs in ontologies and vocabularies; (4) making each IRI provide a knowledge graph with related data about itself; and (5) representing the ontologies and vocabularies themselves as knowledge graphs, typically expressed in RDFS or OWL.[5] The term knowledge graph can also be used to refer to company-internal, proprietary data, whereas LOD emphasizes public sharing of semantic data.

The LOD Cloud has grown to include more than 1200 datasets that adhere to these principles[6], with a total of around 150 trillion edges[7]. Much-used datasets such as DBpedia [166], GeoNames[8], LinkedGeoData [167], and Wikidata[9] link the knowledge graphs in the LOD cloud by offering standard names (IRIs) for people, organizations, places, works, and so on.

8.1 Methodologies and Key Techniques

8.1.1 *Ontologies and vocabularies*

Knowledge graphs and LOD rely heavily on ontologies and vocabularies that offer standard names, in the form of IRIs, for important emergency-related individuals, relationships, and classes. Although numerous general and specific ontologies have been proposed (e.g., [168], [169], [170], [171], [172], [173], [174], [175], [176], [177]), there is currently no consensus. For instance, [178] reviews 26 relevant ontologies, vocabularies, and taxonomies divided into 11 subject areas (disasters, resources, processes, people, organizations, damage, infrastructure, geography, hydrology, meteorology, and topography). Only 14 are represented as knowledge graphs (i.e., using RDF, RDFS, or OWL) and, of those, only 12 are openly available: MOAC (about disasters, see below), FOAF (people and their connections), BIO (biographies), InteLLEO and Organisation Ontology (both about organizations), OTN (infrastructures), GeoNames (geography), Ordnance Survey Hydrology Ontology (hydrology), NNEW (meteorology), USGS CEGIS (topography), Ordnance Survey Building and Places Ontology (buildings and places), E-response Building Pathology Ontology and E-response Building Internal Layout Ontology (both about buildings), and AktiveSA (multi-domain). In [178], it is concluded that while a single, all-encompassing ontology may be neither feasible nor desirable, further work is needed to complete the partial ontologies and make them interoperable.

1. https://www.w3.org/TR/rdf-primer/
2. https://www.w3.org/TR/sparql11-overview/
3. http://www.w3.org/TR/rdf-schema/
4. http://www.w3.org/TR/owl2-overview/

5. The difference is that ontologies tend to be more precise, formal and expressed in OWL; whereas vocabularies are expressed in RDFS.
6. http://lod-cloud.net/
7. http://lodstats.aksw.org/
8. http://www.geonames.org/
9. https://www.wikidata.org/wiki/Wikidata:Introduction

The XML-based Emergency Data Exchange Language (EDXL) [179] facilitates the sharing of emergency information between government agencies and other involved organizations. EDXL is divided into several packages that offer concepts for alerts, information about events, affected areas, and additional image or audio resources (the common alerting protocol package); requesting, responding to, and committing resources (the resource messaging package); field observations, causality, illness, and management reporting (the situation reporting package); hospitals, their statuses, bed capacities, facilities, resources, and services (the hospital availability exchange package); emergency patients (the tracking emergency patients package); high-level information modeling (the reference information model package); and routing XML messages (the distribution elements package). EDXL-RESCUER [168] makes this XML-based standard available as an OWL ontology.

The Management of a Crisis (MOAC) [169] vocabulary links crisis information from three different sources: (a) volunteer and technical committees (the Ushahidi Platform), (b) disaster affected communities, and (c) traditional humanitarian agencies. MOAC is divided into three sections that offer concepts for emergency types, security incidents, and affected populations (the emergency management section); shelters, water, sanitation, food, health, logistics, and telecommunications (the emergency cluster section); and who/what/where/when, needs, and responses (the who-what-where section).[10]

The Humanitarian eXchange Language (HXL) standard aims to improve information sharing during humanitarian crises without creating new reporting burdens[11]. Developed in parallel with a semantic vocabulary [170], HXL defines hashtags for describing places (geolocations, populated places, and administrative units in countries); people and households (affected populations, their needs and characteristics); responses and other operations (relief organizations and their capacities and operations); crises, incidents, and events (crises and events, their causes, impacts and severity); and general metadata (data provenance, approvals, and timestamps). The HXL standard is part of a broader infrastructure that also comprises training, tools, and other materials.

8.1.2 Semantic lifting and linking

Semantic lifting is the process of interpreting and re-representing non-semantic data as knowledge graphs using appropriate ontologies and vocabularies. For example, the HXL infrastructure includes an online tool for annotating spreadsheet data with tags from the HXL vocabulary [170]. Natural language techniques can be used to extract meaning from text through pipelines that include character decoding, tokenization, normalization, stopword removal, stemming/lemmatization, part-of-speech (POS) tagging, and dependency analysis [119]. Recent approaches represent the semantics of words as vectors [180], [181] and analyse them using deep neural networks [182], [183] to support tasks such as text classification, document ranking, reading comprehension, and question answering more precisely.

Microtexts such as Twitter messages and Facebook updates are particularly challenging for natural language techniques because each message is short and offers limited input to analyze, contains little context beyond its associated metadata (particularly if the the message has been sampled from a larger pool), and uses condensed, informal language with abbreviations, slang, misspellings, limited punctuation, and incomplete sentences [184]. Although they are not yet widely used, there exist social media platforms that natively expose their data as knowledge graphs, either through semantic annotations or by other means [185]. Additional approaches to semantic lifting deal with pictures, (live or recorded) audio, and (live or recorded) video. For example, in an emergency situation, computer analyses of pictures posted on social media after an earthquake can be used to identify roadblocks and structural damage to buildings.

When data have been lifted into knowledge graphs, their individuals, relationships, and classes will be uniquely named using standard IRIs, making them linkable to (and thus able to be enriched with) data from other knowledge graphs, such as those in the LOD cloud.

8.1.3 Semantic interoperability

Emergency-relevant datasets that have been lifted into knowledge graphs become easily linkable, not only with open reference datasets, but also with one another. This capacity is extremely valuable in emergency management, where diverse information sources must often be made interoperable rapidly to support complex processing, reasoning, and querying in unforeseeable ways. Indeed, information interoperability has been a central driving force behind all the examples we have so far presented in this section.

8.1.4 Querying, processing, and reasoning

Almost all the semantic applications we have reviewed use SPARQL to query and update the knowledge graphs. The graphs themselves can be manipulated in memory, using APIs such as Apache Jena, or stored in databases, either as semantically-wrapped SQL databases or in native knowledge-graph databases, so-called triple stores (OpenLink Virtuoso[12] is a popular example).

Knowledge graphs are amenable to rule-based reasoning and, if the graph conforms to the more restrictive Description Logic (DL) subset of OWL, also to automated classification, subsumption, consistency checking, and relation checking [186]. For example, [187] applies rules to identify people who have been impacted by earthquakes and adverse weather conditions, whereas [188] uses rules to classify and determine the severity of fire incident reports. Finally, [174] illustrates how an inference engine can be used to determine the emergency level and activate appropriate emergency plans in civil aviation emergencies. Distributed queries and automated reasoning over knowledge graphs constructed from big data are important topics for future research.

10. http://www.observedchange.com/moac/ns/
11. http://hxlstandard.org/
12. https://virtuoso.openlinksw.com/

8.1.5 Big knowledge graphs

The preceding sections have described how knowledge graphs and semantic technologies are being used to handle sources of big data such as remote and human sensors, social media, information services, open data, and government data [189]. As the volume, velocity, and variety of the data grow, the need for semantics that make sense of these data will grow stronger [190] and pose new technical challenges. The term *big linked data* has been coined to reflect this trend [191]. Although mainstream triple stores can reportedly store graphs with more than one trillion edges [192], big linked data will eventually have to be supported by distributed big-data platforms. Graph-oriented big-data technologies such as Google's Pregel [193] and Apache Giraph[13] offer interesting research challenges in this direction.

While techniques for big semantic data analysis can be applied to emergency management out of the box, they may be even more powerful in domain-specific forms [119]. For example, machine learning techniques for natural language processing, ontology learning, and ontology-based learning could be used to tailor information classification and extraction to particular emergency domains. One resulting research challenge is to investigate how domain- and even single crisis-specific techniques, e.g., for social message filtering, could be reused across domains and crises. Another research challenge is to understand how the sub-symbolic outputs of machine learning techniques can best be used in combination with open and interpretable knowledge graphs.

8.2 Real-world Application and Case Study

CERISE-SG: Motivated by the permanent risk of flooding in the Netherlands, the CERISE-SG project [194] explores the use of semantic technologies and LOD to streamline the exchange of data between a regional water board that oversees flood protection and an electricity grid operator. Data types include water levels, which are monitored automatically in real time, the locations of flood-protective assets such as pumping stations, locations of electrical grid assets, status reports from pumping stations and electrical assets, and the affected areas and consumers. Interoperability is facilitated by a bespoke domain ontology that unifies all the data as knowledge graphs. Among the benefits are the flexibility of the underlying data structure, ease of updating data dynamically and in a decentralized manner, external compliance to simplify the link with other data sources, and precise semantics to enable formal reasoning and inference. The authors conclude that semantic technologies and LOD, particularly when they become more widely used and supported by a dedicated crisis ontology, will be highly useful.

Sahana Asia: Built on top of the Sahana platform [195], Sahana Asia [176] uses knowledge graphs to integrate data from external sources such as live earthquake alerts and weather data, known vulnerabilities in cities, open street maps, other map information, and contextual information from open sources such as DBpedia and GeoNames. The data are stored as knowledge graphs using a bespoke disaster data management vocabulary. Sahana Asia can be used to identify potentially disastrous situations in advance, predict vulnerable places and infrastructures based on historical data, visualize data in maps, generate evaluation plans, and disseminate appropriately user-tailored alerts and warnings. The system combines semantic reasoning with conventional analysis approaches such as regression.

Safety Check: Safety Check [187] is an experimental semantic web application that identifies the people affected by a natural or human-made disaster. The application uses a bespoke ontology to integrate personal data from Facebook's Graph API with earthquake and weather alters from public sources and open geographical background information about cities, their geo-coordinates, populations, and areas. The resulting knowledge graph is used to automate reasoning about the affected areas and people and to provide those affected with safety instructions.

9 Conclusion and Future Directions

BDEM is an interdisciplinary and emerging research field that requires collaboration between researchers from various backgrounds. In this paper, we have presented a comprehensive overview of the key concepts of BDEM and discussed the existing technologies and real-world applications from the perspective of interdisciplinary fields. We end this article with a list of open issues and future directions of study:

Reliable and Efficient Disaster-relief Architecture: First, building a high-reliability disaster-relief architecture that handles the sensing and collection of disaster data efficiently in terms of energy and delivery delay remains an open problem. The trade-off between detection accuracy and response complexity should be emphasized.

Accurate and Multidimensional Positioning: Currently, due to the limitation of the infrastructure, GPS data still contains strong noise. During a disaster, the accuracy of location information will be greatly related to the efficiency of rescue. For example, during a fire, the accurate information of the indoor location (location and storey) of victims is crucial for rescue strategies making. Therefore, smartphone-based approaches that can accurately and multidimensionally track human mobility and recognize human behavior in complex environments are urgently needed to satisfy the requirements of disaster-relief applications. Heterogeneous data fusion technologies will be a future direction of research to solve this challenge.

IoT in emergency: During a disaster, people expect to understand the disaster situation from the collected data timely. Thus, quickly gathering the required amount of training data samples is a critical problem. With the development of IoT, large-volume sensing data will be transmitted to aggregation gateways which can solve this problem. However, how to ensure efficient data delivery in DTN-based ECNs in the emergency situation and how to quickly utilize these data under the IoT architecture will be challenging.

Deployment of Computing Resources: The collected data need to be processed quickly to retrieve meaningful information for evacuation and rescue. For this purpose, several groups of edge servers could be employed in a

13. http://giraph.apache.org

decentralized manner in real time. Consequently, the deployment of data processing/computing resources while coordinating with data-collecting devices is another problem that requires further research.

Privacy: Privacy is an important issue in human mobile sensing. To achieve situational awareness, data need to be collected from both public and private sensor networks, as well as smartphone-based applications. This results in privacy issues. Existing social media are used with an annotation functionality, through which private information such as a user's home address, daily office routine, and social activities could be easily inferred from multimedia data including images, audio records, and videos posted on their social media networks. To preserve users' privacy, social media usually allow the privacy level to be tuned when sharing information online. This leads to a trade-off between the privacy level and situational awareness performance: strict privacy controls would limit the useful information that could extracted about disaster scenes.

Emotion Sensing: *Emotion sensing* [196] is an emerging topic in smartphone-based data analytics. Under disaster scenarios, determining the emotions of victims would be useful in providing emotional care so as to help them overcome difficulties and recover from disasters. Although some recent studies [197], [198] have conducted sentiment analysis based on Twitter datasets, it is still challenging to estimate the psychological status of people with high accuracy in the context of disasters. This is an interesting open problem for data-driven emergency management.

Knowledge Transfer: There are typically many reports of major events and disasters (e.g., earthquakes, typhoons, and tsunamis). Seeing and reading about these events may cause others to wonder, "What if this happened in my city?" The same question also intrigues city managers, city planners, and emergency response agencies. Obviously, learning lessons and gaining experience from emergency events can serve as critical examples for designing other cities and ensuring more effective emergency management planning. Thus, models or approaches that transfer human disaster behavior and activities from one disaster to another, and from the affected city to other unaffected cities, are a promising direction for study.

Data-driven relief management optimization: Currently, most of the existing big data-related works on emergency management are focusing on collecting, sensing and predicting the accurate information during emergencies. In the aspect of relief management optimization research, such as emergency logistics optimization, supply-demand matching, most of the works engaged in developing advanced mathematical models but few in designing data-driven methods. Bridging the gaps between the achievements of the two aspects of studies to establish practical data-driven decision support systems will be a significant future direction of study.

Smart Cities with Integrated Disaster-Relief Infrastructures: Finally, establishing smart cities with resilient disaster recovery capabilities is a promising direction for future sustainable development. To achieve this goal, an integrated disaster-relief infrastructure equipped with multiple heterogeneous technologies, including satellite communication networks, aerial-vehicle-based networks, and ground smart-device networks, should be developed. For this, efficient distributed algorithms that can work in decentralized and autonomous environments must be able to coordinate with the integrated infrastructure.

Acknowledgments

This work is supported by the Transnational Partnership for Excellent Research and Education in Big Data and Emergency Management Project funded by the Research Council of Norway (RCN) , the Norwegian Centre for International Cooperation in Education (SiU) through the INTPART programme, and the Grant-in-Aid for Early-Career Scientists (19K15260) of the Japan Ministry of Education, Culture, Sports, Science, and Technology (MEXT).

References

[1] Y. K and H. M, "An example of using its in disasters," in *IPSJ Digital Practice*. IPSJ, 2012.

[2] T. Li, N. Xie, C. Zeng, W. Zhou, L. Zheng, Y. Jiang, Y. Yang, H.-Y. Ha, W. Xue, Y. Huang *et al.*, "Data-driven techniques in disaster information management," *ACM Computing Surveys (CSUR)*, vol. 50, no. 1, p. 1, 2017.

[3] J. Wang, Y. Wu, N. Yen, S. Guo, and Z. Cheng, "Big data analytics for emergency communication networks: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 3, pp. 1758–1778, 2016.

[4] J. Qadir, A. Ali, R. ur Rasool, A. Zwitter, A. Sathiaseelan, and J. Crowcroft, "Crisis analytics: big data-driven crisis response," *Journal of International Humanitarian Action*, vol. 1, no. 1, p. 12, 2016.

[5] V. Hristidis, S.-C. Chen, T. Li, S. Luis, and Y. Deng, "Survey of data management and analysis in disaster situations," *Journal of Systems and Software*, vol. 83, no. 10, pp. 1701–1714, 2010.

[6] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban computing: concepts, methodologies, and applications," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 5, no. 3, p. 38, 2014.

[7] T. Schempp, M. Hong, H. Zhang, R. Akerkar, and A. Schmidt, "An integrated crowdsourced framework for disaster relief distribution," in *2018 5th International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*. IEEE, 2018, pp. 1–4.

[8] T. Xia, X. Song, H. Zhang, X. Song, H. Kanasugi, and R. Shibasaki, "Measuring spatio-temporal accessibility to emergency medical services through big gps data," *Health & place*, vol. 56, pp. 53–62, 2019.

[9] C.-C. Liu, M.-C. Shieh, M.-S. Ke, and K.-H. Wang, "Flood prevention and emergency response system powered by google earth engine," *Remote Sensing*, vol. 10, no. 8, p. 1283, 2018.

[10] Y.-L. Lai, Y.-H. Chou, and L.-C. Chang, "An intelligent iot emergency vehicle warning system using rfid and wi-fi technologies for emergency medical services," *Technology and health care*, vol. 26, no. 1, pp. 43–55, 2018.

[11] Y. Nemoto and K. Hamaguchi, "Resilient ICT research based on lessons learned from the Great East Japan Earthquake," *IEEE Communications Magazine*, vol. 52, no. 3, pp. 38–43, 2014.

[12] J. M. of Internal Affairs and Communications, "Part i – information and communication status at the east japan great earthquake," in *FY2011 White paper for ICT*. MIC, 2012.

[13] F. Harary, "Graph theory," 1969.

[14] H. Frank and I. T. Frisch, "Communication, transmission, and transportation networks," 1971.

[15] A. Itai and M. Rodeh, "The multi-tree approach to reliability in distributed networks," *Information and Computation*, vol. 79, pp. 43–59, 1988.

[16] Q. She, X. Huang, and J. Jue, "How reliable can two-path protection be?" *IEEE/ACM Trans. Networking*, vol. 18, no. 3, pp. 922–933, 2010.

[17] M. Mdard, S. G. Finn, R. A. Barry, and R. G. Gallager, "Redundant trees for preplanned recovery in arbitrary vertex-redundant or edge-redundant graphs," *IEEE/ACM Trans. Networking*, vol. 7, pp. 641–652, 1999.

[18] G. Xue, L. Chen, and K. Thulasiraman, "Quality-of-service and quality-of-protection issues in preplanned recovery schemes using redundant trees," *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS*, vol. 21, no. 8, pp. 1332–1345, 2003.

[19] B. Hao, J. Tang, and G. Xue, "Fault-tolerant relay node placement in wireless sensor networks: formulation and approximation," in *IEEE HPSR?2004*. IEEE, 2004.

[20] W. Zhang, G. Xue, and S. Misra, "Fault-tolerant relay node placement in wireless sensor networks: problems and algorithms," in *IEEE INFOCOM?2007*. IEEE, 2007.

[21] A. Uchiyama, K. Kitada, and T. Higashino, "A scheduling algorithm of cell zooming for energy efficiency in disasters," in *Proceedings of the 13th ACM International Symposium on Mobility Management and Wireless Access*. ACM, 2015, pp. 63–68.

[22] M. Bansal, J. Mehlman, S. Katti, and P. Levis, "Openradio: a programmable wireless dataplane," in *Proceedings of the first workshop on Hot topics in software defined networks*. ACM, 2012, pp. 109–114.

[23] A. Gudipati, D. Perry, L. E. Li, and S. Katti, "Softran: Software defined radio access network," in *Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking*. ACM, 2013, pp. 25–30.

[24] H. Huang, P. Li, S. Guo, and W. Zhuang, "Software-defined wireless mesh networks: architecture and traffic orchestration," *IEEE network*, vol. 29, no. 4, pp. 24–30, 2015.

[25] Q. T. Minh, K. Nguyen, C. Borcea, and S. Yamada, "On-the-fly establishment of multihop wireless access networks for disaster recovery," *IEEE Communications Magazine*, vol. 52, no. 10, pp. 60–66, 2014.

[26] T. Yabe, K. Tsubouchi, A. Sudo, and Y. Sekimoto, "Estimating evacuation hotspots using gps data: What happened after the large earthquakes in kumamoto, japan," in *Proc. of the 5th International Workshop on Urban Computing*, 2016.

[27] M. Krol, Y. Ji, S. Yamada, C. Borcea, L. Zhong, K. Takano *et al.*, "Extending network coverage by using static and mobile relays during natural disasters," in *2016 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA)*. IEEE, 2016, pp. 681–686.

[28] T. Sakano, S. Kotabe, T. Komukai, T. Kumagai, Y. Shimizu, A. Takahara, T. Ngo, Z. M. Fadlullah, H. Nishiyama, and N. Kato, "Bringing movable and deployable networks to disaster areas: development and field test of mdru," *IEEE Network*, vol. 30, no. 1, pp. 86–91, 2016.

[29] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*. ACM, 2005, pp. 252–259.

[30] L. Zhong, K. Garlichs, S. Yamada, K. Takano, and Y. Ji, "Mission planning for uav-based opportunistic disaster recovery networks," in *Wireless Communications and Networking Conference Workshops (WCNCW), 2018 IEEE*. IEEE, 2018.

[31] L. Zhong, Y. Ji, X. Wang, S. Yamada, K. Takano, and G. Xue, "Population-aware relay placement for wireless multi-hop based network disaster recovery," in *Global Communications Conference (GLOBECOM), 2017 IEEE*. IEEE, 2017.

[32] L. Zhong, K. Takano, F. Jiang, X. Wang, Y. Ji, and S. Yamada, "Spatio-temporal data-driven analysis of mobile network availability during natural disasters," in *International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*. IEEE, 2016.

[33] F.-Z. Jiang, L. Zhong, K. Thilakarathna, A. Seneviratne, K. Takano, S. Yamada, and Y. Ji, "Supercharging crowd dynamics estimation in disasters via spatio-temporal deep neural network," in *Data Science and Advanced Analytics (DSAA), 2017 IEEE International Conference on*. IEEE, 2017.

[34] Wikipedia, "Emergency communications network," Accessed 2017-12-27. [Online]. Available: https://en.wikipedia.org/wiki/Emergency_communi\\cations_network

[35] M. Narang, W. Liu, J. Gutierrez, and L. Chiaraviglio, "A cyber physical buses-and-drones mobile edge infrastructure for large scale disaster emergency communications," in *2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)*. IEEE, 2017, pp. 53–60.

[36] T. Higashino, H. Yamaguchi, A. Hiromori, A. Uchiyama, and K. Yasumoto, "Edge computing and iot based research for building safe smart cities resistant to disasters," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*, 2017, pp. 1729–1737.

[37] A. Davis, J. Parikh, and W. E. Weihl, "Edgecomputing: extending enterprise applications to the edge of the internet," in *Proceedings of the 13th international World Wide Web conference on Alternate track papers & posters*. ACM, 2004, pp. 180–187.

[38] S. M. George, W. Zhou, H. Chenji, M. Won, Y. O. Lee, A. Pazarloglou, R. Stoleru, and P. Barooah, "Distressnet: a wireless ad hoc and sensor network architecture for situation management in disaster response," *IEEE Communications Magazine*, vol. 48, no. 3, 2010.

[39] H. Chenji, W. Zhang, R. Stoleru, and C. Arnett, "Distressnet: A disaster response system providing constant availability cloud-like services," *Ad Hoc Networks*, vol. 11, no. 8, pp. 2440–2460, 2013.

[40] A. Fujihara and H. Miwa, "Disaster evacuation guidance using opportunistic communication: The potential for opportunity-based service," in *Big Data and Internet of Things: A Roadmap for Smart Environments*. Springer, 2014, pp. 425–446.

[41] E. M. Trono, Y. Arakawa, M. Tamai, and K. Yasumoto, "Dtn mapex: Disaster area mapping through distributed computing over a delay tolerant network," in *Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU)*. IEEE, 2015, pp. 179–184.

[42] K. Gomez, S. Kandeepan, M. M. Vidal, V. Boussemart, R. Ramos, R. Hermenier, T. Rasheed, L. Goratti, L. Reynaud, D. Grace *et al.*, "Aerial base stations with opportunistic links for next generation emergency communications," *IEEE Communications Magazine*, vol. 54, no. 4, pp. 31–39, 2016.

[43] P. Li, T. Miyazaki, S. Guo, and W. Zhuang, "Online scheduling of mobile stations for disaster management," in *IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–6.

[44] M. Thapa, A. Alsadoon, P. Prasad, L. Pham, and A. Elchouemi, "Impact of using mobile devices in earthquake," in *13th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. IEEE, 2016, pp. 1–6.

[45] P. Li, T. Miyazaki, K. Wang, S. Guo, and W. Zhuang, "Vehicle-assist resilient information and network system for disaster management," *IEEE Transactions on Emerging Topics in Computing*, pp. 1–12, 2017.

[46] A. Valcarce, T. Rasheed, K. Gomez, S. Kandeepan, L. Reynaud, R. Hermenier, A. Munari, M. Mohorcic, M. Smolnikar, and I. Bucaille, "Airborne base stations for emergency and temporary events," in *International Conference on Personal Satellite Services*. Springer, 2013, pp. 13–25.

[47] A. Orsino, L. Militano, G. Araniti, and A. Iera, "Social-aware content delivery with d2d communications support for emergency scenarios in 5g systems," in *Proceedings of 22th European Wireless Conference*. VDE, 2016, pp. 1–6.

[48] C. Meurisch, T. A. B. Nguyen, S. Wullkotte, S. Niemczyk, F. Kohnhäuser, and M. Mühlhäuser, "Nicer911: Ad-hoc communication and emergency services using networking smartphones and wireless home routers," in *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing (Mobihoc)*, 2017, pp. 33:1–33:2.

[49] T. Higuchi, S. Fujii, H. Yamaguchi, and T. Higashino, "Mobile node localization focusing on stop-and-go behavior of indoor pedestrians," *IEEE Transactions on Mobile Computing*, vol. 13, no. 7, pp. 1564–1578, 2014.

[50] Y. Noh, H. Yamaguchi, and U. Lee, "Infrastructure-free collaborative indoor positioning scheme for time-critical team operations," *IEEE Transactions on Systems, Man, and Cybernetics Systems*, 2016.

[51] S. Kojima, A. Uchiyama, M. Shirakawa, A. Hiromori, H. Yamaguchi, and T. Higashino, "Crowd and event detection by fusion of camera images and micro blogs," in *IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, 2017, pp. 213–218.

[52] P. Zuo, Y. Hua, X. Liu, D. Feng, W. Xia, S. Cao, J. Wu, Y. Sun, and Y. Guo, "Bees: Bandwidth-and energy-efficient image sharing for real-time situation awareness," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2017, pp. 1510–1520.

[53] T. Dao, K. Khalil, A. K. Roy-Chowdhury, S. V. Krishnamurthy, and L. Kaplan, "Energy efficient object detection in camera sensor networks," in *2017 IEEE 37th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 2017, pp. 1208–1218.

[54] T. Higuchi, H. Yamaguchi, and T. Higashino, "Mobile devices as an infrastructure: A survey of opportunistic sensing technology," *Journal of information processing*, vol. 23, no. 2, pp. 94–104, 2015.

[55] J. Bradley, J. Barbier, and D. Handler, "Embracing the internet of everything to capture your share of $14.4 trillion," *White Paper, Cisco*, 2013.

[56] F. Bonomi, R. Milito, J. Zhu, and S. Addepalli, "Fog computing and its role in the internet of things," in *Proceedings of the first edition of the MCC workshop on Mobile cloud computing*. ACM, 2012, pp. 13–16.

[57] D. Peng, X. Zhao, Q. Zhao, and Y. Yu, "Smartphone based public participant emergency rescue information platform for earthquake zone "e-explorer"," in *Vibroengineering Procedia*, vol. 5, 2015, pp. 436–439.

[58] R. Han, X. Zhao, Y. Yu, Q. Guan, D. Peng, M. Li, and J. Ou, "Emergency communication and quick seismic damage investigation based on smartphone," *Advances in Materials Science and Engineering*, vol. 2016, 2016.

[59] T. Miyazaki, K. Anazawa, Y. Igarashi, P. Li, and S. Guo, "Resilient information management system for disaster situations," in *IEEE International Conference on Information Networking (ICOIN)*, 2017, pp. 1–6.

[60] K. E. Joyce, S. E. Belliss, S. V. Samsonov, S. J. McNeill, and P. J. Glassey, "A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters," *Progress in Physical Geography*, vol. 33, no. 2, pp. 183–207, 2009.

[61] D. B. A. K. J. Hodgson, M. E., "Geospatial techniques in urban hazard and disaster analysis," p. 327?354, 2010.

[62] L. Wang, K. Lu, P. Liu, R. Ranjan, and L. Chen, "Ik-svd: Dictionary learning for spatial big data via incremental atom update," *Computing in Science Engineering*, vol. 16, no. 4, pp. 41–52, July 2014.

[63] A. Talwalkar, S. Kumar, M. Mohri, and H. Rowley, "Large-scale svd and manifold learning," *Journal of Machine Learning Research*, vol. 14, pp. 3129–3152, 2013. [Online]. Available: http://jmlr.org/papers/v14/talwalkar13a.html

[64] H. Zhang, X. Song, X. Song, D. Huang, N. Xu, R. Shibasaki, and Y. Liang, "Ex-ante online risk assessment for building emergency evacuation through multimedia data," *PloS one*, vol. 14, no. 4, p. e0215149, 2019.

[65] L. Barrington, S. Ghosh, M. Greene, S. Har-Noy, J. Berger, S. Gill, A. Lin, and C. Huyck, "Crowdsourcing earthquake damage assessment using remote sensing imagery," *Annals of Geophysics*, vol. 54, no. 6, 2012.

[66] T. S. S. G. M. Zook, M. Graham, "Volunteered geographic information and crowdsourcing disaster relief: a case study of the haitian earthquake," *World Medical & Health Policy*, vol. 2, no. 2, p. 6?32, 2010.

[67] E. Christenson, M. Elliott, O. Banerjee, L. Hamrick, and J. Bartram, "Climate-related hazards: A method for global assessment of urban and rural population exposure to cyclones, droughts, and floods," *International Journal of Environmental Research and Public Health*, vol. 11, no. 2, pp. 2169–2192, 2014. [Online]. Available: http://www.mdpi.com/1660-4601/11/2/2169

[68] B. R. L.-M. P. F. H. S. K. Fluet-Chouinard, E.; Lehner, "Development of a global inundation map at high spatial resolution from topographic downscaling of coarse-scale remote sensing data," *Remote Sensing of Environment*, vol. 158, pp. 348–361, 2015.

[69] I. T. K. J. W. W. B. B. D. M. K. A. A. A. J. C. A. A. Jason K. Blackburn, Ulrica Diamond and J. J. Glenn Morris, "Household-level spatiotemporal patterns of incidence of cholera, haiti, 2011," *Emerging Infectious Diseases*, vol. 20, no. 9.

[70] B. Burkett and A. Curtis, "Classifying wildfire risk at the building scale in the wildland-urban interface: Applying spatial video approaches to los angeles county," *Risk, Hazards & Crisis in Public Policy*, vol. 2, no. 4, pp. 1–20, 2011.

[71] Z. Zhu, C. E. Woodcock, C. Holden, and Z. Yang, "Generating synthetic landsat images based on all available landsat data: Predicting landsat surface reflectance at any given time," *Remote Sensing of Environment*, vol. 162, no. Supplement C, pp. 67 – 83, 2015.

[72] X. Cao, J. Han, S. Yang, D. Tao, and L. Jiao, "Band selection and evaluation with spatial information," *International Journal of Remote Sensing*, vol. 37, no. 19, pp. 4501–4520, 2016.

[73] A. G. Kalmikov and P. Heimbach, "A hessian-based method for uncertainty quantification in global ocean state estimation," *SIAM Journal on Scientific Computing*, vol. 36, no. 5, pp. S267–S295, 2014.

[74] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[75] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.

[76] X. Lu, L. Bengtsson, and P. Holme, "Predictability of population displacement after the 2010 haiti earthquake," *Proceedings of the National Academy of Sciences*, vol. 109, no. 29, pp. 11 576–11 581, 2012.

[77] J. P. Bagrow, D. Wang, and A.-L. Barabasi, "Collective response of human populations to large-scale emergencies," *PloS one*, vol. 6, no. 3, p. e17680, 2011.

[78] C. Song, T. Koren, P. Wang, and A.-L. Barabási, "Modelling the scaling properties of human mobility," *Nature Physics*, vol. 6, no. 10, pp. 818–823, 2010.

[79] N. Eagle, A. S. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data," *Proceedings of the National Academy of Sciences*, vol. 106, no. 36, pp. 15 274–15 278, 2009.

[80] Z. Fan, X. Song, R. Shibasaki, and R. Adachi, "Citymomentum: an online approach for crowd behavior prediction at a citywide level," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2015, pp. 559–569.

[81] X. Song, Q. Zhang, Y. Sekimoto, R. Shibasaki, N. J. Yuan, and X. Xie, "Prediction and simulation of human mobility following natural disasters," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 8, no. 2, p. 29, 2016.

[82] L. Wang, Z. Yu, B. Guo, T. Ku, and F. Yi, "Moving destination prediction using sparse dataset: A mobility gradient descent approach," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 11, no. 3, p. 37, 2017.

[83] Z. Yang, D. Lian, N. J. Yuan, X. Xie, Y. Rui, and T. Zhou, "Indigenization of urban mobility," *Physica A: Statistical Mechanics and its Applications*, vol. 469, pp. 232–243, 2017.

[84] J. Zhang, B. Guo, Q. Han, Y. Ouyang, and Z. Yu, "Crowdstory: multi-layered event storyline generation with mobile crowdsourced data," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 2016, pp. 237–240.

[85] T. Konishi, M. Maruyama, K. Tsubouchi, and M. Shimosaka, "Cityprophet: city-scale irregularity prediction using transit app logs," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2016, pp. 752–757.

[86] B. Du, C. Liu, W. Zhou, Z. Hou, and H. Xiong, "Catch me if you can: Detecting pickpocket suspects from large-scale transit records," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016, pp. 87–96.

[87] L. Chen, D. Zhang, G. Pan, X. Ma, D. Yang, K. Kushlev, W. Zhang, and S. Li, "Bike sharing station placement leveraging heterogeneous urban open data," in *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2015, pp. 571–575.

[88] X. Song, Q. Zhang, Y. Sekimoto, and R. Shibasaki, "Prediction of human emergency behavior and their mobility following large-scale disaster," in *Proc. of the 20th ACM SIGKDD Conference on Knowledge discovery and data mining*. ACM, 2014, pp. 5–14.

[89] X. Song, Q. Zhang, Y. Sekimoto, T. Horanont, S. Ueyama, and R. Shibasaki, "Intelligent system for human behavior analysis and reasoning following large-scale disasters," *IEEE Intelligent Systems*, vol. 28, no. 4, pp. 35–42, 2013.

[90] X. Song, R. Shibasaki, N. J. Yuan, X. Xie, T. Li, and R. Adachi, "Deepmob: Learning deep knowledge of human emergency behavior and mobility from big and heterogeneous data," *ACM Transactions on Information Systems (TOIS)*, vol. 35, no. 4, p. 41, 2017.

[91] X. Song, Q. Zhang, Y. Sekimoto, T. Horanont, S. Ueyama, and R. Shibasaki, "Modeling and probabilistic reasoning of population evacuation during large-scale disaster," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2013, pp. 1231–1239.

[92] T. Yabe, K. Tsubouchi, A. Sudo, and Y. Sekimoto, "Predicting irregular individual movement following frequent mid-level disasters using location data from smartphones," in *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2016, p. 54.

[93] J. Zhou, H. Pei, and H. Wu, "Early warning of human crowds based on query data from baidu maps: Analysis based on shanghai stampede," in *Big Data Support of Urban Planning and Management*. Springer, 2018, pp. 19–41.

[94] Z. Fan, X. Song, and R. Shibasaki, "Cityspectrum: a non-negative tensor factorization approach," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 2014, pp. 213–223.

[95] D. Lian, C. Zhao, X. Xie, G. Sun, E. Chen, and Y. Rui, "Geomf: joint geographical modeling and matrix factorization for point-of-interest recommendation," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2014, pp. 831–840.

[96] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with gps history data," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 1029–1038.

[97] P. Castro, D. Zhang, and S. Li, "Urban traffic modelling and prediction using large scale taxi gps traces," *Pervasive Computing*, pp. 57–72, 2012.

[98] P.-T. Chen, F. Chen, and Z. Qian, "Road traffic congestion monitoring in social media with hinge-loss markov random fields," in *Data Mining (ICDM), 2014 IEEE International Conference on*. IEEE, 2014, pp. 80–89.

[99] J. Zheng and L. M. Ni, "An unsupervised framework for sensing individual and cluster behavior patterns from human mobile data," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 2012, pp. 153–162.

[100] H. Su, K. Zheng, H. Wang, J. Huang, and X. Zhou, "Calibrating trajectory data for similarity-based analysis," in *Proceedings of the 2013 ACM SIGMOD International Conference on Management of Data*. ACM, 2013, pp. 833–844.

[101] X. Song, H. Kanasugi, and R. Shibasaki, "Deeptransport: Prediction and simulation of human mobility and transportation mode at a citywide level." in *IJCAI*, 2016, pp. 2618–2624.

[102] J. Zhang, Y. Zheng, and D. Qi, "Deep spatio-temporal residual networks for citywide crowd flows prediction." in *AAAI*, 2017, pp. 1655–1661.

[103] M. X. Hoang, Y. Zheng, and A. K. Singh, "Fccf: forecasting citywide crowd flows based on big data," in *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. ACM, 2016, p. 6.

[104] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning," *IEEE Transactions on Intelligent Transportation Systems*.

[105] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015.

[106] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.

[107] X. Ma, H. Yu, Y. Wang, and Y. Wang, "Large-scale transportation network congestion evolution prediction using deep learning theory," *PloS one*, vol. 10, no. 3, p. e0119044, 2015.

[108] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, "Multimodal deep learning," in *Proceedings of the 28th International Conference on Machine Learning (ICML-11)*, 2011, pp. 689–696.

[109] L. Ge, J. Gao, X. Li, and A. Zhang, "Multi-source deep learning for information trustworthiness estimation," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2013, pp. 766–774.

[110] N. Srivastava and R. R. Salakhutdinov, "Multimodal learning with deep boltzmann machines," in *Advances in neural information processing systems*, 2012, pp. 2222–2230.

[111] P. Karuna, M. Rana, and H. Purohit, "Citizenhelper: A streaming analytics system to mine citizen and web data for humanitarian organizations." in *ICWSM*, 2017, pp. 729–730.

[112] J. Preece, *Online Communities: Designing Usability and Supporting Socialbilty*, 1st ed. New York, NY, USA: John Wiley & Sons, Inc., 2000.

[113] P. R. Center, "Social media fact sheet," http://www.pewinternet.org/fact-sheet/social-media, Accessed 2017-09-10.

[114] A. Sheth, "Citizen sensing, social signals, and enriching human experience," *IEEE Internet Computing*, vol. 13, no. 4, pp. 87–92, July 2009.

[115] L. Palen, K. M. Anderson, G. Mark, J. Martin, D. Sicker, M. Palmer, and D. Grunwald, "A vision for technology-mediated support for public participation & assistance in mass emergencies & disasters," in *Proceedings of the 2010 ACM-BCS Visions of Computer Science Conference*, ser. ACM-BCS '10. Swinton, UK, UK: British Computer Society, 2010, pp. 8:1–8:12. [Online]. Available: http://dl.acm.org/citation.cfm?id=1811182.1811194

[116] M. Imran, C. Castillo, F. Diaz, and S. Vieweg, "Processing social media messages in mass emergency: A survey," *ACM Computing Surveys (CSUR)*, vol. 47, no. 4, p. 67, 2015.

[117] H. Purohit, C. Castillo, F. Diaz, A. Sheth, and P. Meier, "Emergency-relief coordination on social media: Automatically matching resource requests and offers," *First Monday*, vol. 19, no. 1, 2013.

[118] K. Glasgow, J. Vitak, Y. Tausczik, and C. Fink, "?with your help... we begin to heal?: Social media expressions of gratitude in the aftermath of disaster," in *Social, Cultural, and Behavioral Modeling: 9th International Conference, SBP-BRiMS 2016, Washington, DC, USA, June 28-July 1, 2016, Proceedings 9*. Springer, 2016, pp. 226–236.

[119] C. Castillo, *Big crisis data: social media in disasters and time-critical situations*. Cambridge University Press, 2016.

[120] J. Howe, "The rise of crowdsourcing," *Wired magazine*, vol. 14, no. 6, pp. 1–4, 2006.

[121] D. C. Brabham, "Crowdsourcing as a model for problem solving: An introduction and cases," *Convergence*, vol. 14, no. 1, pp. 75–90, 2008.

[122] A. Doan, R. Ramakrishnan, and A. Y. Halevy, "Crowdsourcing systems on the world-wide web," *Communications of the ACM*, vol. 54, no. 4, pp. 86–96, 2011.

[123] M. Imran, C. Castillo, J. Lucas, P. Meier, and S. Vieweg, "Aidr: Artificial intelligence for disaster response," in *Proceedings of the 23rd International Conference on World Wide Web*. ACM, 2014, pp. 159–162.

[124] P. Meier, *Digital humanitarians: how big data is changing the face of humanitarian response*. Crc Press, 2015.

[125] N. Geographic, "Ushahidi haiti project," http://voices.nationalgeographic.org/2012/07/02/crisis-mapping-haiti/uhp2/, Accessed 2017-09-10.

[126] T. H. Nazer, G. Xue, Y. Ji, and H. Liu, "Intelligent disaster response via social media analysis a survey," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 46–59, 2017.

[127] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: real-time event detection by social sensors," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 851–860.

[128] J. Yin, A. Lampert, M. Cameron, B. Robinson, and R. Power, "Using social media to enhance emergency situation awareness," *IEEE Intelligent Systems*, vol. 27, no. 6, pp. 52–59, 2012.

[129] A. Olteanu, C. Castillo, F. Diaz, and S. Vieweg, "Crisislex: A lexicon for collecting and filtering microblogged communications in crises." in *ICWSM*, 2014.

[130] I. P. Temnikova, C. Castillo, and S. Vieweg, "Emterms 1.0: A terminological resource for crisis tweets." in *ISCRAM*, 2015.

[131] S. E. Vieweg, "Situational awareness in mass emergency: A behavioral and linguistic analysis of microblogged communications," Ph.D. dissertation, University of Colorado at Boulder, 2012.

[132] I. Varga, M. Sano, K. Torisawa, C. Hashimoto, K. Ohtake, T. Kawai, J.-H. Oh, and S. De Saeger, "Aid is out there: Looking for help from tweets during a large scale disaster." in *ACL (1)*, 2013, pp. 1619–1629.

[133] S. Ranganath, S. Wang, X. Hu, J. Tang, and H. Liu, "Facilitating time critical information seeking in social media," *IEEE Transactions on Knowledge and Data Engineering*, 2017.

[134] C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in *Proceedings of the 20th international conference on World wide web*. ACM, 2011, pp. 675–684.

[135] A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, "Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy," in *Proceedings of the 22nd international conference on World Wide Web*. ACM, 2013, pp. 729–736.

[136] K. Starbird, J. Maddock, M. Orand, P. Achterman, and R. M. Mason, "Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing," *iConference 2014 Proceedings*, 2014.

[137] A. Olteanu, S. Vieweg, and C. Castillo, "What to expect when the unexpected happens: Social media communications across crises," in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 2015, pp. 994–1009.

[138] K. Starbird, G. Muzny, and L. Palen, "Learning from the crowd: collaborative filtering techniques for identifying on-the-ground twitterers during mass disruptions," in *Proceedings of 9th International Conference on Information Systems for Crisis Response and Management, ISCRAM*, 2012.

[139] H. Purohit, S. Bhatt, A. Hampton, V. L. Shalin, A. P. Sheth, and J. M. Flach, "With whom to coordinate, why and how in adhoc social media communications during crisis response," in *Proceedings of 11th International Conference on Information Systems for Crisis Response and Management, ISCRAM*, 2014, pp. 787–791.

[140] S. Kumar, F. Morstatter, R. Zafarani, and H. Liu, "Whom should i follow?: identifying relevant users during crises," in *Proceedings of the 24th ACM conference on hypertext and social media*. ACM, 2013, pp. 139–147.

[141] C. Reuter, O. Heger, and V. Pipek, "Combining real and virtual volunteers through social media." in *ISCRAM*, 2013.

[142] A. Opdyke and A. Javernick-Will, "Building coordination capacity: Post-disaster organizational twitter networks," in *IEEE Global Humanitarian Technology Conference (GHTC 2014)*, 2014, pp. 86–92.

[143] H. Purohit and J. Chan, "Classifying user types on social media to who-what-where coordination during crisis response," in *Proceedings of 14th International Conference on Information Systems for Crisis Response and Management, ISCRAM*, 2017, pp. 656–665.

[144] L. Palen, K. Starbird, S. Vieweg, and A. Hughes, "Twitter-based information distribution during the 2009 red river valley flood threat," *Bulletin of the Association for Information Science and Technology*, vol. 36, no. 5, pp. 13–17, 2010.

[145] C. Hui, Y. Tyshchuk, W. A. Wallace, M. Magdon-Ismail, and M. Goldberg, "Information cascades in social media in response to a crisis: a preliminary model and a case study," in *Proceedings of the 21st International Conference on World Wide Web*. ACM, 2012, pp. 653–656.

[146] X. Lu and C. Brelsford, "Network structure and community evolution on twitter: human behavior change in response to the 2011 japanese earthquake and tsunami," *Scientific reports*, vol. 4, 2014.

[147] H. Purohit, Y. Ruan, D. Fuhry, S. Parthasarathy, and A. P. Sheth, "On understanding the divergence of online social group discussion." 2014.

[148] W. Rand, J. Herrmann, B. Schein, and N. Vodopivec, "An agent-based model of urgent diffusion in social media," *Journal of Artificial Societies and Social Simulation*, vol. 18, no. 2, p. 1, 2015.

[149] E. Serrano, C. A. Iglesias, and M. Garijo, "A survey of twitter rumor spreading simulations," in *Computational Collective Intelligence*. Springer, 2015, pp. 113–122.

[150] J. Lingad, S. Karimi, and J. Yin, "Location extraction from disaster-related microblogs," in *Proceedings of the 22nd international conference on world wide web*. ACM, 2013, pp. 1017–1020.

[151] S. E. Middleton, L. Middleton, and S. Modafferi, "Real-time crisis mapping of natural disasters using social media," *IEEE Intelligent Systems*, vol. 29, no. 2, pp. 9–17, 2014.

[152] A. Crooks, A. Croitoru, A. Stefanidis, and J. Radzikowski, "# earthquake: Twitter as a distributed sensor system," *Transactions in GIS*, vol. 17, no. 1, pp. 124–147, 2013.

[153] J. Chae, D. Thom, Y. Jang, S. Kim, T. Ertl, and D. S. Ebert, "Public behavior response analysis in disaster events utilizing visual analytics of microblog data," *Computers & Graphics*, vol. 38, pp. 51–60, 2014.

[154] J. Rogstadius, M. Vukovic, C. Teixeira, V. Kostakos, E. Karapanos, and J. A. Laredo, "Crisistracker: Crowdsourced social media curation for disaster awareness," *IBM Journal of Research and Development*, vol. 57, no. 5, pp. 4–1, 2013.

[155] A. Sheth, A. Jadhav, P. Kapanipathi, C. Lu, H. Purohit, G. A. Smith, and W. Wang, "Twitris: A system for collective social intelligence," in *Encyclopedia of Social Network Analysis and Mining*. Springer, 2014, pp. 2240–2253.

[156] S. Kumar, G. Barbier, M. A. Abbasi, and H. Liu, "Tweettracker: An analysis tool for humanitarian and disaster relief." in *ICWSM*, 2011.

[157] P. Meier, "Crisis mapping in action: How open source software and global volunteer networks are changing the world, one map at a time," *Journal of Map & Geography Libraries*, vol. 8, no. 2, pp. 89–100, 2012.

[158] T. Schempp, H. Zhang, A. Schmidt, M. Hong, and R. Akerkar, "A framework to integrate social media and authoritative data for disaster relief detection and distribution optimization," *International Journal of Disaster Risk Reduction*, p. 101143, 2019.

[159] H. Purohit, M. Dalal, P. Singh, B. Nissima, V. Moorthy, A. Vemuri, V. Krishnan, R. Khursheed, S. Balachandran, H. Kushwah *et al.*, "Empowering crisis response-led citizen communities: Lessons learned from jkfloodrelief. org initiative," in *Strategic Management and Leadership for Systems Development in Virtual Spaces*. IGI Global, 2016, pp. 270–292.

[160] P. Neis and D. Zielstra, "Recent developments and future trends in volunteered geographic information research: The case of openstreetmap," *Future Internet*, vol. 6, no. 1, pp. 76–106, 2014.

[161] T. Berners-Lee, J. Hendler, O. Lassila *et al.*, "The semantic web," *Scientific american*, vol. 284, no. 5, pp. 28–37, 2001.

[162] N. Shadbolt, T. Berners-Lee, and W. Hall, "The semantic web revisited," *IEEE intelligent systems*, vol. 21, no. 3, pp. 96–101, 2006.

[163] G. Babitski, S. Bergweiler, O. Grebner, D. Oberle, H. Paulheim, and F. Probst, "Soknos–using semantic technologies in disaster management software," in *Extended Semantic Web Conference*. Springer, 2011, pp. 183–197.

[164] S. Potter, Y. Kalfoglou, H. Alani, M. Bachler, S. Buckingham Shum, R. Carvalho, A. Chakravarthy, S. Chalmers, S. Chapman, B. Hu *et al.*, "The application of advanced knowledge technologies for emergency response," 2007.

[165] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data-the story so far," *Semantic services, interoperability and web applications: emerging concepts*, pp. 205–227, 2009.

[166] J. Lehmann, R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P. N. Mendes, S. Hellmann, M. Morsey, P. Van Kleef, S. Auer *et al.*, "Dbpedia–a large-scale, multilingual knowledge base extracted from wikipedia," *Semantic Web*, vol. 6, no. 2, pp. 167–195, 2015.

[167] C. Stadler, J. Lehmann, K. Höffner, and S. Auer, "Linkedgeodata: A core for a web of spatial open data," *Semantic Web*, vol. 3, no. 4, pp. 333–354, 2012.

[168] R. Barros, P. Kislansky, L. Salvador, R. Almeida, M. Breyer, and L. G. Pedraza, "Edxl-rescuer ontology: Conceptual model for semantic integration." in *Proc. Int'l Conf. Information Systems for Crisis Response and Management (ISCRAM)*, 2015.

[169] M. Limbu, D. Wang, T. Kauppinen, and J. Ortmann, "Management of a crisis (moac) vocabulary specification," Tech. Rep., 2012.

[170] C. Keßler and C. Hendrix, "The humanitarian exchange language: Coordinating disaster response with semantic web technologies," *Semantic Web*, vol. 6, no. 1, pp. 5–21, 2015.

[171] I. S. M. Iwanaga, T.-M. Nguyen, T. Kawamura, H. Nakagawa, Y. Tahara, and A. Ohsuga, "Building an earthquake evacuation ontology from twitter," in *Granular Computing (GrC), 2011 IEEE International Conference on*. IEEE, 2011, pp. 306–311.

[172] P. D. Haghighi, F. Burstein, A. Zaslavsky, and P. Arbon, "Development and evaluation of ontology for intelligent decision support in medical emergency management for mass gatherings," *Decision Support Systems*, vol. 54, no. 2, pp. 1192–1204, 2013.

[173] M. K. A. Hassan and Y.-H. Chen-Burger, "A communication and tracking ontology for mobile systems in the event of a large scale disaster," in *Agent and Multi-Agent Systems: Technology and Applications*. Springer, 2016, pp. 119–137.

[174] H. Wang, Z. Pan, and W. Zhou, "The design and implementation of information system in civil airport emergency management based on semantic web," in *Control and Decision Conference (CCDC), 2013 25th Chinese*. IEEE, 2013, pp. 2392–2396.

[175] S. Consoli, A. Gangemi, A. G. Nuzzolese, S. Peroni, D. R. Recupero, and D. Spampinato, "Setting the course of emergency vehicle routing using geolinked open data for the municipality of catania," in *European Semantic Web Conference*. Springer, 2014, pp. 42–53.

[176] T. Silva, V. Wuwongse, and H. N. Sharma, "Disaster mitigation and preparedness using linked open data," *Journal of Ambient*

Intelligence and Humanized Computing, vol. 4, no. 5, pp. 591–602, 2013.

[177] N. García-Santa, E. García-Cuesta, and B. Villazón-Terrazas, "Controlling and monitoring crisis," in *International Semantic Web Conference*. Springer, 2015, pp. 46–50.

[178] S. Liu, C. Brewster, and D. Shaw, "Ontologies for crisis management: A review of state of the art in ontology design and usability." in *Proc. Int'l Conf. Information Systems for Crisis Response and Management (ISCRAM)*, 2013.

[179] O. E. M. T. Committee *et al.*, "Emergency data exchange language (edxl) distribution element, v. 1.0," *OASIS Standard EDXL-DE v1*, 2006.

[180] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," in *Advances in neural information processing systems*, 2013, pp. 3111–3119.

[181] J. Pennington, R. Socher, and C. Manning, "Glove: Global vectors for word representation," in *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 2014, pp. 1532–1543.

[182] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of deep bidirectional transformers for language understanding," *arXiv preprint arXiv:1810.04805*, 2018.

[183] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. Salakhutdinov, and Q. V. Le, "Xlnet: Generalized autoregressive pretraining for language understanding," *arXiv preprint arXiv:1906.08237*, 2019.

[184] A. Sheth and K. Thirunarayan, "Semantics empowered web 3.0: managing enterprise, social, sensor, and cloud-based data and services for advanced applications," *Synthesis Lectures on Data Management*, vol. 4, no. 6, pp. 1–175, 2012.

[185] S. Tramp, P. Frischmuth, N. Arndt, T. Ermilov, and S. Auer, "Weaving a distributed, semantic social network for mobile users," in *Extended Semantic Web Conference*. Springer, 2011, pp. 200–214.

[186] D. Nardi and R. J. Brachman, "An introduction to description logics," in *The Description Logic Handbook: Theory, Implementation and Applications*, F. Baader, Ed. Cambridge university press, 2003, ch. 1.

[187] Y. Pandey and S. Bansal, "Safety check: a semantic web application for emergency management," in *Proceedings of The International Workshop on Semantic Big Data*. ACM, 2017, p. 7.

[188] A. Schulz and H. Paulheim, "Combining government and linked open data in emergency management," 2012.

[189] C. Bizer, P. Boncz, M. L. Brodie, and O. Erling, "The meaningful use of big data: four perspectives–four challenges," *ACM SIGMOD Record*, vol. 40, no. 4, pp. 56–60, 2012.

[190] K. Janowicz, F. Van Harmelen, J. A. Hendler, and P. Hitzler, "Why the data train needs semantic rails," *AI Magazine*, 2014.

[191] B. Hu, N. Carvalho, L. Laera, and T. Matsutsuka, "Towards big linked data: a large-scale, distributed semantic data storage," in *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services*. ACM, 2012, pp. 167–176.

[192] Oracle, "Oracle spatial and graph: Benchmarking a trillion edges RDF graph," Oracle White Paper, November 2016. [Online]. Available: http://download.oracle.com/otndocs/tech/semantic_web/pdf/OracleSpatialGraph_RDFgraph_1_trillion_Benchmark.pdf

[193] G. Malewicz, M. H. Austern, A. J. Bik, J. C. Dehnert, I. Horn, N. Leiser, and G. Czajkowski, "Pregel: a system for large-scale graph processing," in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data*. ACM, 2010, pp. 135–146.

[194] R. Roller, J. Roes, and E. Verbree, "Benefits of linked data for interoperability during crisis management," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, no. 3, p. 211, 2015.

[195] M. Careem, C. De Silva, R. De Silva, L. Raschid, and S. Weerawarana, "Sahana: Overview of a disaster management system," in *Information and Automation, 2006. ICIA 2006. International Conference on*. IEEE, 2006, pp. 361–366.

[196] M. Chen, Y. Zhang, Y. Li, S. Mao, and V. C. Leung, "Emc: Emotion-aware mobile cloud computing in 5g," *IEEE Network*, vol. 29, no. 2, pp. 32–38, 2015.

[197] A. Montejo-Ráez, E. Martínez-Cámara, M. T. Martín-Valdivia, and L. A. Ureña-López, "Ranked wordnet graph for sentiment polarity classification in twitter," *Computer Speech & Language*, vol. 28, no. 1, pp. 93–107, 2014.

[198] L. Y.-F. Su, M. A. Cacciatore, X. Liang, D. Brossard, D. A. Scheufele, and M. A. Xenos, "Analyzing public sentiments online: Combining human-and computer-based content analysis," *Information, Communication & Society*, vol. 20, no. 3, pp. 406–427, 2017.

Xuan Song is an Excellent Young Researcher of Japan MEXT, and an Associate Professor at The University of Tokyo. He received the Ph.D. degree in signal and information processing from Peking University, China, in 2010. His main research interest are AI and its related research areas, such as data mining, intelligent system, computer vision, and robotics, especially on intelligent surveillance and information system design, mobility and spatio-temporal data mining, sensor fusion, and machine learning algorithms development. By now, he has published more than 60 technical publications in journals, book chapter, and international conference proceedings, including more than 30 high-impact papers in top-tier publications for computer science and robotics, such as ACM TOIS, ACM TIST, IEEE T-PAMI, KDD, UbiComp, IJCAI, AAAI, ICCV, CVPR, ECCV, ICRA and etc.

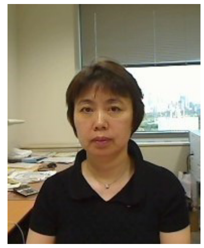
Haoran Zhang received the Ph.D. degree in oil-gas storage and transportation engineering in 2018 from China University of Petroleum-Beijing. After graduation, he works as a researcher at The University of Tokyo. His current research interests include system optimization and data mining. He has published about 60 SCI papers, and 10 conference papers in many top international journals and conferences, such as Applied Energy, Energy, Journal of Cleaner Production and IEEE Transactions on Industrial Informatics.

Rajendra Akerkar is a Professor and Head of Big Data Research Group at Western Norway Research Institute. He has 28 years of experience in knowledge representation and reasoning, ubiquitous data analytics and requirement engineering. He is actively involved in several international research innovation initiatives for more than 22 years. Currently he coordinates three major research initiatives funded by European Commission (Horizon 2020 programme) and the Research Council of Norway on two different topics: ubiquitous data-driven mobility/transport and Big Data in emergency management. He is a member of IEEE Task Force on Disaster Resilient Smart World. He is serving as an Associate Editor of International Journal of Metadata, Semantics and Ontologies, and Knowledge Management Track Editor of Web Intelligence, an international journal.

Huawei Huang (M'16) received his Ph.D in Computer Science and Engineering from the University of Aizu, Japan, in 2016. He is currently an Associate Professor at School of Data and Computer Science with Sun Yat-Sen University, China. His research interests mainly include SDN/NFV, edge computing, and blockchain system. He received the best paper award of TrustCom-2016. He used to be a visiting scholar at the Hong Kong Polytechnic University (2017-2018), a post-doctoral research fellow of JSPS (2016-2018), an Assistant Professor with Kyoto University, Japan (2018-2019). He is a member of the IEEE.

Song Guo (M02SM11) is a Full Professor with the Department of Computing, Hong Kong Polytechnic University, Hong Kong. He received his PhD degree in computer science from University of Ottawa. He has authored or coauthored over 450 papers in major conferences and journals. His current research interests include big data, cloud and edge computing, mobile computing, and distributed systems. Prof. Guo was a recipient of the 2019 TCBD Best Conference Paper Award, the 2018 IEEE TCGCC Best Magazine Paper Award, the 2017 IEEE SYSTEMS JOURNAL Annual Best Paper Award, and six other Best Paper Awards from IEEE/ACM conferences. He was an IEEE Communications Society Distinguished Lecturer. He has served as an Associate Editor of IEEE TPDS, IEEE TCC, IEEE TETC, etc. He also served as the general and program chair for numerous IEEE conferences. He currently serves in the Board of Governors of the IEEE Communications Society.

Lei Zhong received his Ph. D. degree on informatics from The Graduate University for Advances Studies in Japan. From 2011 to 2017, he involved in many national research projects related with wireless communication and networking when he worked in National Institute of Information and Communications Technology (NICT), National Institute of Informatics (NII) and The University of Tokyo respectively.He currently is a senior researcher and project manager in Toyota Motor Corporation, where he is responsible for the research and architect of future network and computing infrastructure for connected vehicles. He is also serving as a Toyota delegate in several related standardization organizations such as 3GPP and AECC. His research interests include edge computing, vehicular networking, software-defined networking, Internet-of-Things and Big Data. He has been published more than 30 technical papers and applied several patents in these fields.



Yusheng Ji (M'94-SM18) received the B.E., M.E., and D.E. degrees in electrical engineering from the University of Tokyo. She joined the National Center for Science Information Systems (NACSIS), Japan, in 1990. She is currently a Professor with the National Institute of Informatics (NII), and SOKENDAI (the Graduate University for Advanced Studies), Japan. Her research interests include network architecture, resource management in wireless networks, and mobile computing. She has served as a Board Member of Trustees of IEICE, Steering Committee Member of Quality Aware Internet SIG, Expert Member of IEICE Technical Committees on Communication Quality, Symposium Co-Chair of IEEE GLOBECOM in 2012 and 2014, Track Chair of IEEE VTC 2016 Fall and 2017 Fall, Associate Editor of IEICE Transactions and IPSJ Journal. She is an Expert Member of IEICE Technical Committees on Internet Architecture, Steering Committee Member of Internet and Operation Technologies SIG of IPSJ, Editor of IEEE Transactions of Vehicular Technology, Symposium Co-Chair of IEEE ICC2020, and TPC Member of IEEE INFOCOM, ICC, GLOBECOM, WCNC, etc.

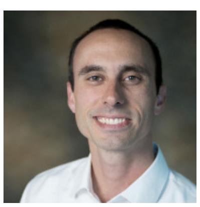
Andreas L. Opdahl is Professor of Information Systems Development at the University of Bergen, Norway. He received his Ph.D. from the Norwegian University of Science and Technology in 1992. His research interests include ontologies and knowledge graphs, enterprise and IS modelling, as well as safety and security requirements. Opdahl is the author, co-author or co-editor of more than a hundred peer-reviewed research papers that have been cited more than 3500 times. He is a member of IFIP WG5.8 on Enterprise Interoperability and WG8.1 on Design and Evaluation of Information Systems. He serves regularly as a reviewer for premier international journals such as MISQ, JAIS, and ISR and as an organiser of the most renowned international conferences and workshops in his fields of interest.

Hemant Purohit is an Assistant Professor of Information Sciences and Technology in the Volgenau School of Engineering at George Mason University, Fairfax, USA. His research interests include social web mining, semantic computing, and human-AI collaboration with applications in disaster informatics and smart city services. Contact him at hpurohit@gmu.edu.

Andr e Skupin is a Professor of Geography and Co-Director of the Center for Information Convergence and Strategy (CICS) at San Diego State University. He received a Dipl.-Ing. Degree in Cartography from the Technical University Dresden, Germany, and a PhD in Geography at the University at Buffalo, United States. His research expertise is centered on data visualization, machine learning and knowledge management, with applications in such diverse domains as biomedicine, demography, and criminology.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBDATA.2020.2972871, IEEE Transactions on Big Data

Akshay Pottathil is a strategist with extensive experience in developing public-private partnerships. He has a proven ability to conduct accurate needs analysis, assess technical capabilities, conduct strategic planning, build and motivate teams, and manage high-risk/high-reward projects. At San Diego State University, he was honored with the Faculty of the Year award in the Homeland Security program and has been granted multiple Presidential Leadership Fund awards. His efforts have spanned across North America, Europe, the Middle East, and South East Asia. Pottathil serves as Co-Director of San Diego State Universitys Center for Information Convergence and Strategy (CICS), Co-Director of the Center for Data Analytics and Intelligence (CENDAI) at the University of Life Sciences Prague, in addition to other academic and non-academic appointments.



Aron Culotta is an associate professor of computer science at the Illinois Institute of Technology in Chicago, where he leads the Text Analysis in the Public Interest lab. His research focuses on extracting socially valuable insights from online social networks. He is a former Microsoft Live Labs Fellow with a PhD in computer science from the University of Massachusetts, Amherst. His work has received best paper awards at the Association for the Advancement of Artificial Intelligence (AAAI) conference and the Conference on Computer-Supported Social Work and Social Computing (CSCW).