Enhancing Emergency Response Systems through Leveraging Crowdsensing and Heterogeneous Data Analytics

Mervat Abu-Elkheir
Faculty of Computer and Information Sciences
Mansoura University
Mansoura, Egypt
mfahmy78@mans.edu.eg

Hossam S. Hassanein
School of Computing
Queen's University
Kingston ON, Canada
hossam@cs.queensu.ca

Sharief M.A. Oteafy
School of Computing
Queen's University
Kingston ON, Canada
oteafy@cs.queensu.ca

***Abstract*— Robust and prompt emergency response is a crucial service that smart cities should provide to citizens, communities, and corporations. Emergency management strategies that are currently supported by cities yield pre-determined protocols that can only handle well-understood incidents. However, there are incidents whose nature, shape, scale, and timing are not as predictable. The lack of adequate data management platforms to harvest emergency-related data from the proliferation of data sources scattered around a city is a major shortfall in current emergency response and risk assessment processes. We propose an improved information infrastructure to assist emergency personnel in responding effectively and proportionally to large-scale, distributed, unstructured natural and man-made hazards such as multi-vehicle accidents, outbreaks of human or animal diseases, major weather events, large fires, and terrorist attacks. The proposed infrastructure will crowdsource the multitude of human and physical sensing resources that can generate data about incidents (e.g. smartphones, sensors, vehicles, etc.) in order to build a comprehensive understanding of emergency situations and provide situational awareness and recommendations to emergency teams on the scene. Our infrastructure consists of three components: (1) large-scale crowdsensing and data quality valuation, (2) heterogeneous data integration and analytics, and (3) decision making, alternative generation and recommendations. Leveraging crowdsensing and heterogeneous data analytics will improve the response coordination to critical incidents and real-time incident management, which will contribute to saving lives and reducing injuries, improving the quality of life, and saving resources by deploying them more effectively.**

Keywords—smart city; emergency response; crowdsourcing

I. Introduction

Urbanization is creating larger cities with larger populations and more complex infrastructure. To make cities better equipped with the means to provide better living conditions, smart city initiatives are being proposed to provide monitoring and maintenance of critical infrastructure, as well as planning and optimization of resource utilization [1]. Two factors help achieve this vision: real-time data collection from physical and virtual sensing devices around the city, and the interconnection of smart devices across systems using network equipment [2].

Emergency management and response is a crucial service that is provided by cities at different scales. Emergency management strategies that are currently supported by cities produce protocols that can only handle well-understood incidents. However, as cities become more complex and as residents' dynamics evolve, the scale, nature, shape, and timing of incidents are becoming harder to predict. Timely production and dissemination of information about incidents are crucial for coordination efforts, mobilization of response resources, situation analysis and evaluation and future prevention and preparedness strategies. This timely provisioning creates an atmosphere of trust and credibility among city residents in highly charged situations, and provides first responders with critical information about incidents and their spatial contexts.

To realize the full potential of real-time data availability and internetworked smart devices across a city, this paper presents an information infrastructure for emergency management and response. The expansion of incident data sources beyond traditional emergency management channels will be tackled through the discovery and valuation of human and physical sensing resources and the integration of crowdsourced, heterogeneous data. The efficient delivery of data and information will be tackled through mission-critical communication networks and hotspots that utilize smartphones, mobiles resources (e.g. drones and vehicles), and wearables for emergency information delivery. The timely analysis of crowdsourced, heterogeneous data with variable levels of accuracy and consistency will be tackled through the development of a robust computational sensemaking and analytics platform. This platform will integrate algorithms to mitigate uncertainty, incorporate the temporal and geospatial nature of data, and produce actionable insights and strategic decision making. The infrastructure enables building a comprehensive understanding of a given incident and provides timely situational awareness and recommendations to emergency teams on the scene.

The remainder of the paper is organized as follows. Section II outlines the smart emergency management infrastructure. In section III, we describe the crowdsensed data collection methods and. Section IV provides details on the robust delivery of data

978-1-5090-0304-4/16/$31.00 ©2016 IEEE

via dynamic provisioning. Section V highlights the methods and techniques involved in sensemaking. In section VI, example response services that can be supported by the proposed infrastructure are described. Finally, section VII concludes the paper.

II. Smart Emergency Management Infrastructure

A high level view of the proposed information infrastructure for emergency management is illustrated in Fig. 1. To address the issues around dynamic, real-time, serious incident response, four elements are presented: i) crowdsourcing emergency data; ii) dynamic provisioning; iii) analytics and sensemaking; and, iv) response services.

Real-time data collection and validation involves techniques to gather heterogeneous data from sources of qualitatively different kinds. The proposed techniques address three barriers to real-time data access: i) degradation at the site of the incident or in transmission; ii) uncertainty due to stress or disinformation; and, iii) varying levels of significance based on context and data veracity. Provisioning techniques will dynamically configure communication capacity between the site(s) of an incident and locations where analytics will be done; and to dynamically provision the storage and computational resources needed to carry out the analytics. Analytics techniques will focus on learning the type, scope, severity and dynamic properties of incidents from heterogeneous data. The predictive analytics will be developed to serve short-term prediction as an incident is progressing, as well as long-term prediction to explore incidents behavior and distinguishing features that can be used for future proactive preparedness measures. Services for response teams and survivors will be designed with focus on supporting immediate search and rescue efforts, such as efficient and accurate navigation and localization in indoor and outdoor crisis zones, evacuation routes and plans, and healthcare monitoring and alerting services.

To highlight the potential of our infrastructure, we depict what would be the cascading responses to an example disaster; the Lac Mégantic fuel explosion in Quebec, in Fig. 2. We highlight some initial responses for the system, and the ensuing processes, actions and analytics that will unfold to address such a complex disaster involving people, structures, and a multiplicity of responders. The details of each element will be explored in the following sections.

III. Crowdsourcing Emergency Data

Incident data can come from a wide range of conventional sensors in proximity to an incident (CCTV cameras, weather stations, radar, emergency room admissions); from information known to and disseminated by the media; from data posted to social media sites; and from important emerging channels like vehicle sensors, vehicular networks and connected Internet of Things (IoT) devices. Data from these sources comes in multiple formats (text message or voice), varies by size (temperature or video) and varies by rate and update interval. This heterogeneous data must either be collected in one place, or processed in a distributed way. Then, the varying timeframes of the data need to be synchronized, and the data itself needs to be unified both for syntax and semantics. To address these issues, we propose two core components: the first will address resource discovery and enlisting, and the second will address the assessment of data quality.

A. Resource Discovery and Enlisting on the Fly

The lack of resource discovery schemes with standardized protocols and resource utility evaluation hinder the adoption of diverse and localized data resources (sensors, cameras, smartphones, vehicles, drones, wearable devices, etc.) for emergency response systems. The most relevant research efforts focus on web-based approaches for resource discovery [3], with resource descriptions being defined in offline mode and with little emphasis on contextual discovery and ranking of resources.

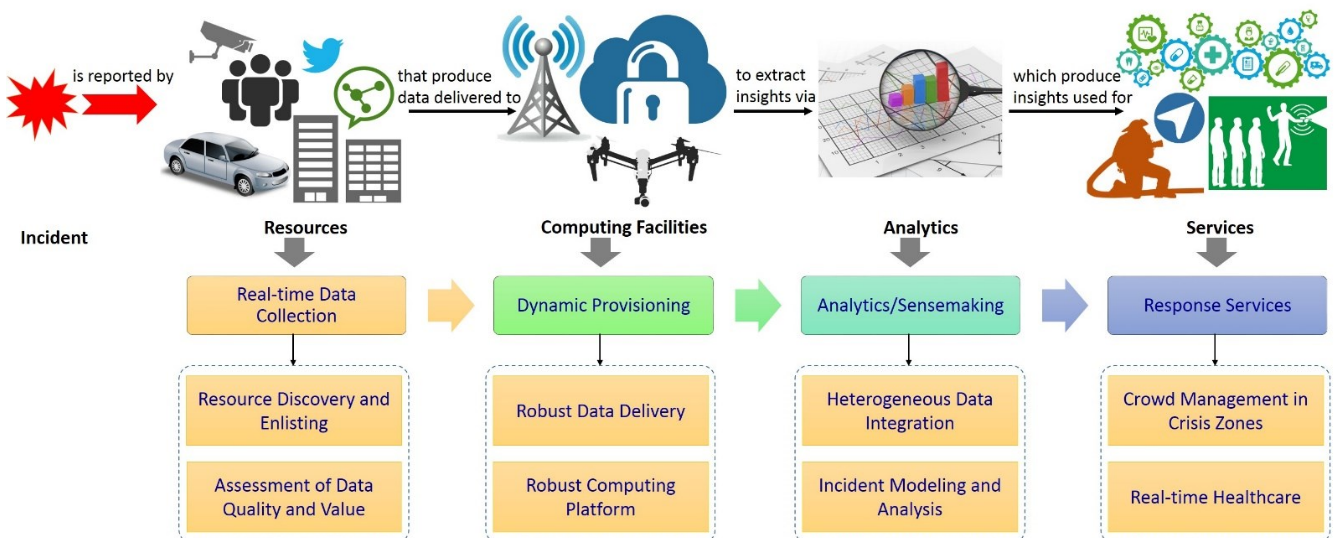

Fig. 1. Emergency management system overview.

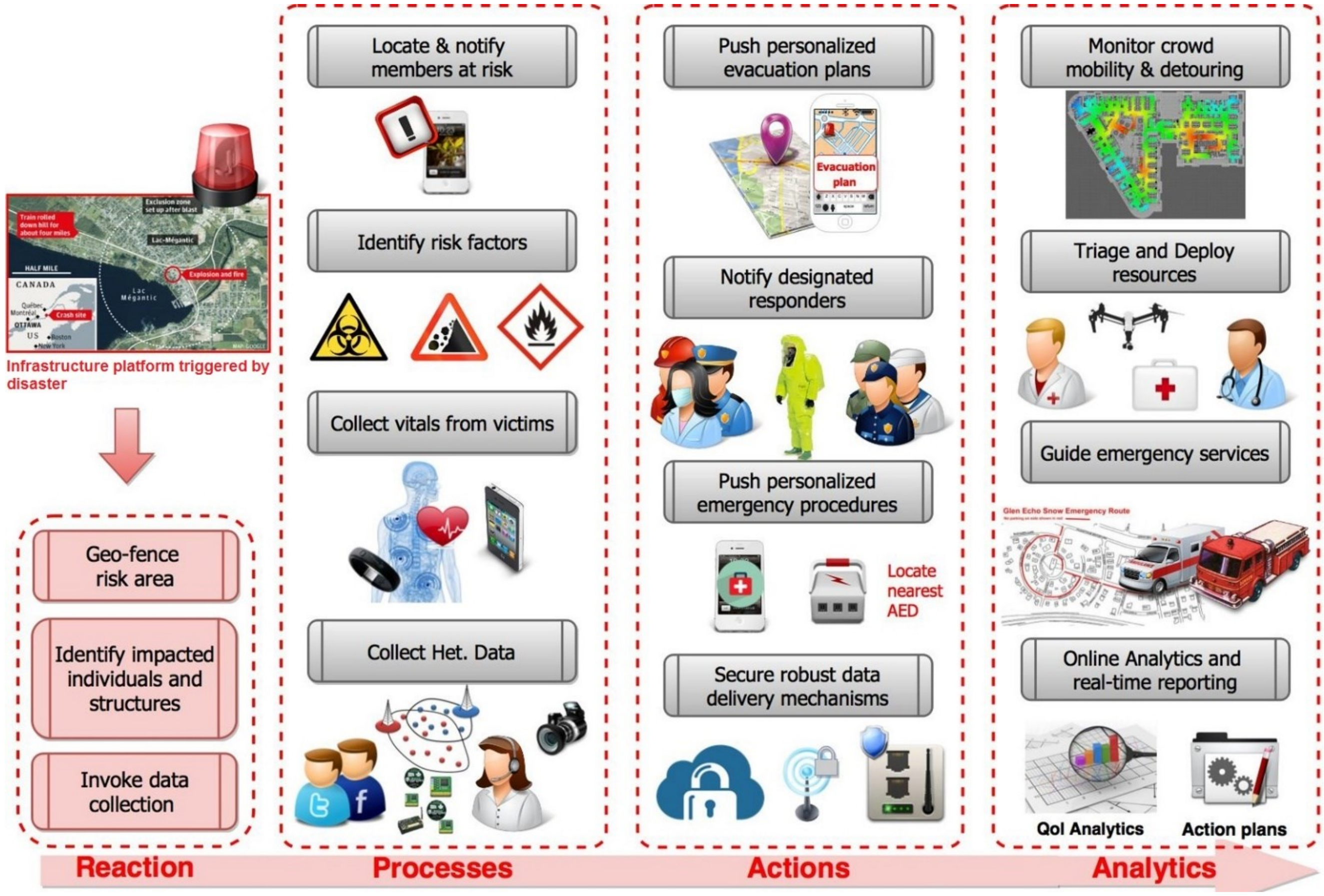

Fig. 2. Emergency management system interactions and workflow scenario.

The resource discovery and enlisting component will identify the availability of resources that are most relevant to the characteristic elements of a given crisis situation, eliciting factors such as: hardware profile, data integrity, functional capabilities, levels of operation, power consumption, location(s), availability and region/time of sensing fidelity. This information will be used to build standardized XML/RDF representations in the form of profiles and service descriptors for each resource. Quantifiable measures will be devised to compute normalized and comparable metrics of the utility of data sources as well as the veracity of the data they produce. These metrics will be used to dynamically index and rank resources based on their extracted attributes in order to enable effective arbitration and resource selection. Dynamic negotiation of functional requirements based on the available resources and their valuation at any given time will be supported to ensure that emergency application running cost is minimized. This component will deploy mechanisms that enable hierarchical resource discovery initiated by crisis management units as well as peer-to-peer resource discovery initiated by peers within an incident zone to discover data sources in their vicinity.

B. *Assessment of the Quality and Value of Data*

Existing sensing systems adopt a collect-and-report model, whereby collected data is indiscriminately pushed onto the networking infrastructure, regardless of the Quality of Data (QoD) or its value (VoD). Aside from the scalability issues produced by this model, establishing reliable response systems is not attainable over inconsistently collected and reported data. Thus, the future of ubiquitous sensing is hampered by the sheer volume of reported data and its uncalibrated discrepancies [4].

The assessment component will deploy a value-attribution model that scales with data abundance and adapts to varying VoD metrics that are governed by the types of crisis situations. The goal is to empower data collection and information fusion on the basis of adaptive QoD and VoD metrics. Using this component, emergency responders will be able to identify the quality and value of data received and data providers will be incentivized to provide higher quality data. Low QoD indices will be used to reduce requests for unwanted data with lower impact to emergency responders and to communicate with local data providers in order to throttle traffic from these sources. Throttling unwanted data with low quality will save precious bandwidth over potentially strained communication networks. A core component of computing QoD will be measures for trust-establishment and adaptation, whereby readings will be cross-referenced with "trusted" resources deployed by emergency responders or designated anchor resources. The assessment component will include a uniform protocol for dictating data aggregation and QoD-based pruning of low-quality input to establish quality attribution to data polled from heterogeneous resources. A scalable value-attribution model for crowd-solicited sensing will be devised to encourage higher-quality reporting of critical data in crises situations. The model will assign value metrics based on the context, scarcity of data, trust profiles of producers and responsiveness to the events and emergency responders (i.e. access latency to resources and their data).

IV. Dynamic Provisioning

When an unexpected incident occurs, there is typically a sharp increase in demand for communications capacity as sensors become fully utilized and as users increase cell phone and wireless network usage to call family members or to connect with social media. Incidents often occur in a confined area, overwhelming available capacity. Known techniques for addressing unexpected demand include microcell set up and carrier capacity-sharing. Unfortunately, neither of these is available quickly. An incident also creates the need for substantial computational capacity at short notice, preferably at the scene. To enable robust communication networks, our infrastructure incorporates the contextual enlisting and deployment of physical mobile resources as ad hoc communication hotspots. We also present a "fog infrastructure" that incorporates cloudlet and cloud facilities for data filtering, fusion and analysis both at the edge to support emergency teams in emergency scenarios and at the backend to support preparedness strategies.

A. *Robust Data Delivery*

The communication network infrastructure has grown to become a heterogeneous wireless networking environment with the Internet as backbone, cellular networks providing large-area coverage, and various application-specific networks such as vehicular networks and wireless sensor networks. However, recent natural disasters demonstrated that this infrastructure cannot meet service demands and provide reliable information delivery in a disastrous situation where the service demand increases beyond the norm [5]. As a result, research initiatives on disaster resilient wireless networking were investigated [6] [7]. Improving network reliability via fault resilient approaches has been mostly limited for wireline backbone networks [8]. The wireless communications tools that are provided to first responders are designed to reliably deliver voice-only communications using narrowband technologies like Terrestrial Trunked Radio [9]. This resulted in a segmented spectrum allocation where different first responders groups use different frequency channels making seamless (voice-only) communication between different teams infeasible and severely limiting data communication.

To provide resilient wireless communication during disasters, the robust data delivery component will include an optimal strategy for deploying movable routers (vehicles, drones, smartphones, etc.) to provide necessary emergence communication service coverage. Hotspot profiling for mission critical communication will be supported via the integration with the resource discovery component. Resources with communication capabilities that are within the vicinity of first responders will be recruited on an as-needed basis, catering for resource-rich and resource-poor scenarios (scenarios where communication links are available and little recruitment is needed versus scenarios where communication links are compromised). Hotspots profiles will be built to identify the availability, signal quality, security, reliability, and coverage of the communications resources available for first responders in a crisis zone. Techniques to effectively distribute data traffic load over limited network resources will be designed to support maximal throughput. Reliable MAC and networking protocols will be designed to minimize the transmission collision probability and delay, while guaranteeing the throughput for mobile hotspots and connected devices on the scene. Network-assisted information dissemination will be integrated with multiple-hop peer-to-peer transmission to guarantee information availability even in the absence of a communication infrastructure. This integration will be bolstered by the development of effective and efficient information diffusion schemes that exploit opportunistic relaying, spatial locations of end users, as well as user mobility.

B. *Robust Computing Platform*

Timely, accurate and effective use of available information is vital to the rapid decision-making required in emergency response scenarios [10]. Cloud computing is a natural approach to providing the computing resources needed by responders in the field to perform real-time analysis of data as an incident develops. It is, however, predicated on reliable, high-bandwidth end-to-end network connections. Incidents such as natural disasters and major accidents are characterized by unpredictable or non-existent network connectivity. By creating a robust computation platform, real-time analytics and effective decision-making in the face of an unpredictable and dynamic environment can be supported.

Fog computing provides a highly virtualized platform that provides computing, storage and network resources between end devices and traditional cloud data centers [11]. Our fog computing platform will have a two-tier structure. The first tier will be a permanent backbone of resources hosted on a public or private cloud. The second tier will be a set of cloudlets that can act as proxies for the first-tier cloud or can function independently [12]. The cloudlets can be positioned prior to an incident or deployed/recruited by first responders on arrival. Once deployed, the cloudlets will support data collection, cleaning and analysis in cooperation with the first tier cloud, or independently in the case of failed connectivity. Spark is a popular framework for in-memory cluster computing. It supports rapid processing by retaining the data in main memory and avoiding writing to disk [13]. It is suited to the quick response times expected in edge analytics for emergency response. This framework will be supported by robust scheduling, synchronization, and recovery algorithms to coordinate processing across cloud-cloudlet tiers in the face of failure.

V. Analytics and Sensemaking

To develop thresholds that trigger particular emergency response actions, it is necessary to understand how incidents develop and evolve. This will be done by computationally modeling a given emergency situation to understand, at an abstract level, what is happening so that the response can be as appropriate and effective as possible. Prediction tasks will be performed to develop an understanding of the following: what kind of incident is happening (determined as early as possible); how serious is it; is it getting worse or better? What does this trajectory look like? The computational modeling of incident situations will leverage data heterogeneity to enable the comprehensive analysis of data from diverse sources.

A. *Heterogeneous Data Integration*

Multisensory-based system can provide data for different event detection tasks including sentiment estimation, object tracking, location estimation, trajectory observation, person detection, person recognition, and activity detection. The performance of these tasks is mainly based on physical sensors and is measured mostly in terms of accuracy [14]. Hence there is a need to fuse crowdsourced data with sensory data and assess the quality of the fused information. Existing research hardly utilized the multi-sensory evidences obtained on-the-fly in modeling data quality attributes.

The data gathered in emergency situations poses challenges on how to efficiently identify events based on 1) fine granular hard sensor measurements and 2) coarse sources of unstructured text snippets and images in social media. It is not practical to fuse and integrate data into rigid schemas before analysis, since the different schemas of data sources cannot be predicted until an incident has occurred and its collocated data sources have been discovered and solicited. Therefore, we propose to perform data fusion and integration dynamically by developing abstractions for each of the data sources based on their types, data generation rates and data formats, without forcing streams from those sources to conform to rigid schemas. The heterogeneous data integration component will include feature identification and alignment mechanisms to dynamically identify, measure and model contextual incident features from multimodal data. The component will also employ probabilistic latent fusion model that unifies sensor data with social media messages in order to optimize real-time event processing and provide an up-to-date situation map. Our goal encompasses a seamless consolidation of sensory data fusion techniques and methodologies whilst the network continues its ordinary operation.

B. *Incident Modeling and Analysis*

The incident modeling and analysis component will focus on producing the most useful possible models to make predictions that can be used by incident commanders in the short term, downstream medical care and logistics in the medium term and recovery in the long term. Sensemaking has two aspects: developing models of the evolving incident and deciding which results of the models should be disseminated to those who make tactical or strategic decisions. Incident modeling has been limited mostly to transportation incidents [15].

Predictive algorithms that will be designed for incident modeling and analysis will use data of unprecedented variety to predict unusual properties such as incident severity, casualty numbers, collateral impacts and economic effects. Furthermore, the production of predictions will need to be made with varying deadlines according to the nature and impact of different incidents. Incident risk models will be designed to be sufficiently concrete that they can be encoded as predictive labels, so that predictions can be made from collected data about incidents. Algorithms will involve spatio-temporal clustering of data about known incidents and create an empirical taxonomy of incidents. Existing regression, deep learning, and topological data analysis techniques will be extended to predict the likelihood and severity of the different types of incidents, based on the risk models developed. An improved Kanri distance approach can be developed for translating the predictive models into action. In order to optimize insight delivery to first responders, game theory will be explored as a tool for sharing information/insights and participating in strategic decision making.

VI. RESPONSE SERVICES

The ultimate goal of an emergency response system is to provide the necessary services that save lives. Applications that can be provided to first responders and survivors will provide information for navigation through emergency zones, recommendation of safe routes, and real-time health services. These information services will be provided across diverse platforms, such as smartphones, drones, vehicles, and wearable technology.

A. *Crowd Management in Crisis Zones*

First responders and citizens/vehicles will need immediate assistance in order to maneuver them out of danger zones to safety. Even though navigation and localization services are pervasively available in today's smartphones and vehicles, they do not provide enough granular accuracy for many crisis situations, especially in urban areas and inside buildings. The infrastructure providing such services may even be compromised in the aftermath of major disasters.

The crowd management component will focus on mechanisms to identify and localize survivors and first responders in an incident zone, as well as mechanisms for zone characterization to facilitate navigation for individuals as well as crowds. Crowdsourced data from sources verified to be in the impact zone of an incident will be collected and integrated at drones deployed to the scene of the incident. Drone technology, which has been used for fleet navigation [16], will be used to provide navigation and localization services on the scene. Drones will feed this data to multiple modules in order to support real-time and continuous navigation and localization of people within the zone. Navigation and localization will be augmented by methodologies to improve occluded-environment tracking as well as zone models that are inferred from zone topology information. Localization will be realized by designing an individual and crowd activity recognition mechanism that is integrated within a generic geo-positioning framework in order to provide effective geo-referencing. Activity recognition will be performed based on heterogeneous sensory streams from other data sources, especially data provided by drones.

The zone's contextual model can be used to identify safe destinations (e.g. exit locations), optimizing the distribution of load on destination points according to their current capacities as well as the crowd size. The crowd management component will also incorporate a pathway recommendation system for pedestrians (first responders and survivors), with recommendations based on safe destinations and environmental tracking. Recommendations will be designed in the form of visual and audio cues that are platform-independent (working on smartphones and wearable devices). Wearable technology promises to change the way that evacuation planning is done

[17]. Zone modeling will also be used in evacuation planning mechanisms for indoor environments.

B. Real-time Healthcare

Wearable devices continuously measure and preprocess the physiological signs from human body and transmit the useful information to the smartphones [18]. Real-time healthcare will support health probing by deploying software services to probe a given crisis zone for potential survivors' physical conditions from wearable/smart monitoring devices and resources. Using wearable technology based health probing will take into consideration the efficient utilization of limited energy and the enhancement of service quality provisioning for vital physiological information transmission. Alert and visualization systems for hazards and diseases will push essential information to citizens in the form of visual environmental hazard and disease spread maps. Insight fusion algorithms from environmental modeling and crowdsourced data reports will be used to predict environmental hazards and assess their potential impact, as well as predict the onset and spread patterns of epidemics.

VII. Conclusions and Future Work

In this paper, we presented an ambitious vision of an information infrastructure for enhancing emergency response systems that incorporates crowdsensing and heterogeneous data analytics. This information infrastructure, when realized, will transform how the current emergency management systems in cities operate and provide life-saving services to citizens. This infrastructure can be integrated with existing risk databases maintained by a city's emergency management centers. We project that the realization of this information infrastructure will improve the response coordination to critical incidents and real-time incident management, which will contribute to saving lives and reducing injuries, improving quality of life, and saving resources by deploying them more effectively.

References

[1] T. Nam and T. Pardo, "Conceptualizing Smart City with Dimensions of Technology, People, and Institutions," in *The Proceedings of the 12th Annual International Conference on Digital Government Research*, 2011.

[2] C. Harrison, B. Eckman, R. Hamilton, P. Hartswick, J. Kalagnanam, J. Paraszczak and P. Williams, "Foundations for Smarter Cities," *IBM Journal of Research and Development,* vol. 54, no. 4, pp. 1-16, 2010.

[3] S. Datta and C. Bonnet, "Search Engine Based Resource Discovery Framework for Internet of Things," in *IEEE 4th Global Conference on Consumer Electronics*, Osaka, Japan, 2015.

[4] R. Ferrando-Llopis, D. Lopez-Berzosa and C. Mulligan, "Advancing value creation and value capture in data-intensive contexts," in *IEEE International Conference on Big Data*, Silicon Valley, CA, 2013.

[5] Y. Ran, "Considerations and suggestions on improvement of communication network disaster countermeasures after the Wenchuan earthquake," *IEEE Communications Magazine,* vol. 49, no. 1, pp. 44-47, 2012.

[6] T. Sakano, Z. Fadlullah, T. Ngo, H. Nishiyama, M. Nakazawa, F. Adachi, N. Kato, A. Takahara, T. Kumagai, H. Kasahara and S. Kurihara, "Disaster-resilient networking: a new vision based on movable and deployable resource units (MDRUs)," *IEEE Network Magazine,* vol. 27, no. 4, pp. 40-46, 2013.

[7] C. Nelson, B. Steckler and J. Stamberger, "The evolution of hastily formed networks for disaster response technologies, case studies, and future trends," in *Proceedings of the IEEE Global Humanitarian Technology Conference*, Seattle, WA, 2011.

[8] S. Trajanovski, F. Kuipers, A. Ilić, J. Crowcroft and P. Van Mieghem, "Finding critical regions and region-disjoint paths in a network," *IEEE/ACM Transactions on Networking,* vol. 23, no. 3, pp. 908-921, 2015.

[9] G. Baldini, S. Karanasios, D. Allen and F. Vergari, "Survey of Wireless Communication Technologies for Public Safety," *IEEE Communications Surveys and Tutorials,* vol. 16, no. 2, pp. 619-641, 2014.

[10] N. Bessis, E. Asimakopoulou, T. French, P. Norrington and F. Xhafa, "The Big Picture, from Grids and Clouds to Crowds: A Data Collective Computational Intelligence Case Proposal for Managing Disasters," in *The International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, 2010.

[11] F. Bonomi, R. Milto, 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*, 2012.

[12] M. Satyanarayanan, P. Bahl, R. Caceres and N. Davies, "The Case for VM-Based Cloudlets in Mobile Computing," *IEEE Pervasive Computing,* vol. 8, no. 4, pp. 14-23, 2009.

[13] M. Zaharia, M. Chowdhury, T. Das, A. Dave, J. Ma, M. McCauley, M. Franklin, S. Shenker and I. Stoica, "Resilient Distributed Datasets: A Fault-Tolerant Abstraction for In-Memory Cluster Computing," in *Proceedings of the 9th USENIX Symposium on Networked Systems Design and Implementation*, 2012.

[14] P. Atrey, M. Hossain, A. El Saddik and M. Kankanhalli., "Multimodal Fusion for Multimedia Analysis: A Survey," *Springer Multimedia Systems Journal,* vol. 16, no. 6, pp. 345-379, 2010.

[15] Y. Ma, R. Fries, M. Chowdhury and I. Inamdar, "Evaluation of the integrated application of intelligent transportation system technologies using stochastic incident generation and resolution modeling," *Simulation,* vol. 88, no. 1, pp. 123-133, 2012.

[16] P. Mosterman, D. Sanabria, E. Bilgin, K. Zhang and J. Zander, "A Heterogeneous Fleet of Vehicles for Automated Humanitarian Missions," *Computing in Science & Engineering,* vol. 16, no. 3, pp. 90-95, 2014.

[17] M. D'Orazio, S. Longhi, P. Olivetti and G. Bernardini, "Design and experimental evaluation of an interactive system for pre-movement time reduction in case of fire," *Automation in Construction,* vol. 52, pp. 16-28, 2015.

[18] I. Cleland, L. Guerrero and J. Bravo, "Big Data Processing Using Wearable Devices for Wellbeing and Healthy Activities Promotion.," in *7th International Work-Conference on Ambient Assisted Living. ICT-based Solutions in Real Life Situations*, Puerto Varas, Chile, 2015.

[19] F. Villanueva, M. Santofimia, D. Villa, J. Barba and J. Lopez, "Civitas: The Smart City Middleware, from Sensors to Big Data," in *Seventh International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS)*, Taichung, 2013.