==
Engineering & Technology in India www.engineeringandtechnologyinindia.com
Vol. 1:5 December 2016
==

Big Data in Disaster Management

Jebas Sinthiya.I, PG Scholar
Shanmugapriya. E.

==

Abstract

Big Data is a frame work to use in particularly disaster prone areas of the globe. It investigates the nature of social media generated during disaster and defines a list of content categories taking into consideration for the information in disaster phases. Smart cities governance can leverage on this Big Data to plan effective disaster management. Disaster management is a crucial and urgent research issue. Emergency Communication Networks (ECNs) provide fundamental functions for disaster management, because communication service is generally unavailable due to large-scale damage and restrictions in communication services .Big data analytics in the disaster area provides possible solutions to understand the situations happening in disaster areas, so that limited resources can be optimally deployed based on the analysis results. In this paper, we study big data in disaster management and its characteristics, challenges, and uses.

Key Words: Disaster Management, Big Data, Volume, Velocity, Emergency Communication Networks.

1. Introduction

Big Data is the name given to our ever-increasing ability to collect more data from a multitude of sources, and analyze it for insights using advanced computer algorithms. Humans can't provide a better understanding of situations and solutions to problems. Disasters are big, messy and noisy situations, and exactly the sort of conditions in which Big Data can help to make sense of the chaos. The massive amounts of data that we are generating with mobile phones, satellites and social media can all play a part in providing clues to the best way to respond to a situation in four phases of disaster management: (Neal,1997) Prevention, Preparedness, Response, Recovery. The DM Act 2005 uses the following definition for disaster: "Disaster" means a catastrophe, mishap, calamity or grave occurrence in any area, arising from natural or manmade causes, or by accident or negligence which results in substantial loss of life or human suffering or damage to, and destruction of, property, or damage to, or degradation of, environment, and is of such a nature or magnitude as to be beyond the coping capacity of the community of the affected area."

The United Nations defines a disaster as a serious disruption of the functioning of a community or a society. Disasters involve widespread human, material, economic or environmental impacts, which exceed the ability of the affected community or society to cope using its own resources. The Red Cross and Red Crescent societies define disaster management as the organization and management of resources and responsibilities for dealing with all humanitarian aspects of emergencies, in particular preparedness, response and recovery in order to lessen the impact of disasters.

The UNISDR defines disaster risk management as the systematic process of using administrative decisions, organization, operational skills and capacities to implement policies, strategies and coping capacities of the society and communities to diminish the impacts of natural hazards and related environmental and technological disasters. This comprises of all forms of activities, including structural and non-structural measures to avoid (prevention) or to limit (mitigation and preparedness) adverse effects of hazards. UNISDR3 has proposed the following definition for the term Disaster Management (UNISDR 2015b): "The organization, planning and application of measures preparing for, responding to and, initial recovery from disasters." As per this definition, 'Disaster Management' focuses on creating and implementing preparedness and others plans to decrease the impact of disasters and build back better. Failure to create/apply a plan could result in damage to life, assets and lost revenue. However, it may not completely avert or eliminate the threats. The term Disaster Management as used in the NPDM

2009 and the DM Act 2005 document is comprehensive covering all aspects – disaster risk reduction, disaster risk management, disaster preparedness, disaster response, and post-disaster recovery. This document uses the term with the same meaning as defined in the DM Act 2005: “A continuous and integrated process of planning, organising, coordinating and implementing measures which are necessary or expedient" for the following: 1) Prevention of danger or threat of any disaster, 2) Mitigation or reduction of risk of any disaster or its severity or consequences, 3) Capacity-building, 4) Preparedness to deal with any disaster, 5) Prompt response to any threatening disaster situation or disaster, 6) Assessing the severity or magnitude of effects of any disaster 7) Evacuation, rescue and relief, and 8) Rehabilitation and reconstruction.”

1.1 Types of Disasters

There is no country that is immune from disaster, though vulnerability to disaster varies. There are four main types of disaster.

- Natural disasters: including floods, hurricanes, earthquakes and volcano eruptions that have immediate impacts on human health and secondary impacts causing further death and suffering from (for example) floods, landslides, fires, tsunamis.
- Environmental emergencies: including technological or industrial accidents, usually involving the production, use or transportation of hazardous material, and occur where these materials are produced, used or transported, and forest fires caused by humans.
- Complex emergencies: involving a break-down of authority, looting and attacks on strategic installations, including conflict situations and war.
- Pandemic emergencies: involving a sudden onset of contagious disease that affects health, disrupts services and businesses, and brings economic and social costs.

Any disaster can interrupt essential services, such as health care, electricity, water, sewage/garbage removal, transportation and communications. The interruption can seriously affect the health, social and economic networks of local communities and countries. Disasters have a major and long-lasting impact on people long after the immediate effect has been mitigated. Poorly planned relief activities can have a significant negative impact not only on the disaster victims but also on donors and relief agencies. So it is important that physical therapists join established programmes rather than attempting individual efforts.

Local, regional, national and international organizations are all involved in mounting a humanitarian response to disasters. Each will have a prepared disaster management plan. These plans cover prevention, preparedness, relief and recovery.

1.2 Levels of Disasters

The disaster management and its planning at various tiers must take into account the vulnerability of disaster-affected area, and the capacity of the authorities to deal with the situation. Using this approach, the High Power Committee on Disaster Management5, in its report of 2001, categorized disaster situations into three 'levels': L1, L2, and L3. The period of normalcy, L0, should be utilized for disaster risk reduction. Level-L1: The level of disaster that can be managed within the capabilities and resources at the District level. However, the state authorities will remain in readiness to provide assistance if needed. Level-L2: This signifies the disaster situations that require assistance and active mobilization of resources at the state level and deployment of state level agencies for disaster management. The central agencies must remain vigilant for immediate deployment if required by the state. Level-L3: This corresponds to a nearly catastrophic situation or a very large-scale disaster that overwhelms the State and District authorities. The categorization of disaster situations into levels L0 to L3 finds no mention in DM Act 2005. Further, the DM Act does not have any provision for notifying any disaster as a ‘national calamity’ or a ‘national disaster’.

1.3 Big Data

The concept of Big Data project is fundamentally related to computer science since the beginning of computing. The term Big Data describes amounts of data obtained with technological means that are normally unusable by humans due to volume and which with appropriate automated processing will extract actionable information. [1]

1.4 Big Data Characteristics

Big Data may be characterized as having four dimensions: Data Volume, measuring the amount of data available, with typical data sets occupying many terabytes. Data velocity is a measure of the rate of data creation, streaming and aggregation. Data variety is a measure of the richness of data representation – text, images, videos etc. Data value, measures the usefulness of data in making decisions. [2]. A further characteristic has recently appeared, namely Variability, which represents the number of changes in the structure of the data their interpretation. Gartner [3] summarizes this in the definition of Big Data as high volume, velocity and variety information assets that demand cost effective processing.

2. Big Data Research Challenges In Disaster Management

The first common challenge reported was the insufficient levels of implementation for each monitored activity. For example, although DRM plans or risk sensitive building codes exist they are not enforced because of a lack of government capacity or public awareness or because so much development takes place in the informal sector. Risk information acquired through assessments is often not translated into policy partly because policy makers are not aware of how to use such information. Staging public awareness raising campaigns, while useful, run the risk of being a one-time event and may not bring any real change in people's behaviour or actions. In other words, it is not sufficient to have risk assessment data and institutional arrangements in place; it is important to consider how these elements actually lead to changes in behavior at all levels in a way that leads to an improved management of risks. A second common challenge highlighted by many countries is the need to strengthen local capacities to implement disaster risk management, including through establishing local level mechanisms and risk assessments. Weak capacity at the local level undermines the implementation of building codes and land use plans. National policies also need to be adapted to the local context (e.g. the national school curricula on DRR that can be tailored to local risks and needs). Small-scale events that many countries struggle with are local in scope.

A third challenge refers to how climate change issues are integrated into DRM (e.g. risk assessment, research, building codes, and land use planning) given that climate change will lead to shifts in risk patterns. Some countries have already combined DRM and climate change adaptation policies and created a common platform to discuss how both need to be mainstreamed into national and local-level policies. While steps have been taken, there is still long way to go before effective policy coordination on climate change and DRM is the norm. Fourth, DRM policymakers have difficulty in obtaining political and economic commitment due to other competing needs and priorities. While many agree that reducing disaster risks is important for saving lives and property, few countries have appropriate measures in place because other issues (e.g. poverty reduction, economic growth, social welfare and education) require greater attention and funding. This has resulted in the insufficient earmarking of financial resources for DRM policies. Land use planners also face difficulty in balancing DRR needs with economic ones. DRM policy makers are in need of clear evidence, including cost-benefit analysis, to convince public and politicians that commitment to DRM is as practical and necessary as any other priority. Another common challenge refers to poor coordination between stakeholders, and a lack of information sharing, including with respect to risk assessment, monitoring and evaluation, early warning, disaster response and other DRM activities.

Mainstreaming DRR in all policy areas and ensuring the commitment of sector agencies is important in preventing new risks from arising and also helps stakeholders address existing risks and strengthen the resiliency of society. Finally, while many countries are still engaged in moving from a response based emergency management paradigm towards the disaster risk reduction paradigm embodied by the HFA, yet others are already pushing the boundaries beyond the HFA towards a new paradigm in which disaster risk management becomes a hallmark of good development

2.1 Quality of Service and Quality of Information in Big Data

For using big data analysis to achieve effective disaster management, the underlying infrastructure must provide high quality of service (QoS). While the QoS requirements may differ for various disaster situations, we outline some examples here as illustration. Given the urgency of the response actions when dealing with most disasters, it is imperative for the infrastructure to provide real-time performance. This includes real-time data analysis to accurately predict the impact of an approaching hazard as well as the best way of effectively responding to the disaster. It also includes real-time communication to ensure that correct data are gathered about the environment, such as the location of people who need help, the best routes for going to a disaster site and for helping people move away from disaster sites. Real-time communication is also needed to ensure that various emergency response teams can coordinate their actions in optimally responding to a disaster.

Given the criticality of disaster response situations, it is also important to ensure that the service will be highly reliable and available in spite of the adverse environmental conditions during such situations, including physical damages, power outages, floods, etc. Hence, the big data storage, analysis, and transmission services must be able to operate in spite of such adverse conditions. Redundancy alone is not adequate since one type of hazard may impact all the redundant units. Hence, this requires the use of diversity in addition to redundancy to ensure high reliability and availability. For example, computing and sensor resources can be deployed at different geographical locations and different communication methods can be used to ensure continuous access to the data.

Given the evolving nature of disasters and disaster response strategies, it is also important to ensure that the big data supporting infrastructure is sufficiently maintainable. This includes methods of ensuring that the infrastructure can be easily upgraded and also to be able to rapidly repair or replace damaged units. Given the sensitivity of some types of data that can substantially help disaster response, such as the location of people and their medical conditions, it is also important to ensure that the service meets high levels of security. This includes high levels of privacy and confidentiality as well as assurance that the information used to guide the response to a disaster are correct and not corrupted. For tele-operation, tele-health, and other remote actions, it is also important to ensure high levels of cyber security. In particular, the infrastructure must ensure that only authorized emergency response personnel can control remote units. Cloud computing platform resources can be leveraged to support big data storage and analysis. Multiple cloud computing platforms at geographically diverse locations can be used to tolerate different hazards. This natural combination of redundancy and diversity can be leveraged to achieve high performance real time big data analysis. This diversification can also be used to achieve highly secure storage and computing by splitting confidential data across multiple sites and using big data analysis methods that can work directly on encrypted data without needing to decrypt the data.

Dependable communication methods are needed that can operate under severe operational conditions, including power outages, damaged communication lines, disruption of wireless signals, damaged communication signal transmission units, etc. Also, it must be ensured that emergency response personnel will be able to access the big data platform and coordinate their actions with other teams. This must be done in spite of areas that may have communication dead-spots, such as underground tunnels, etc. For example, it can require the deployment of sonar, light, and other communication methods in addition to the usual electrical signals for communication. Usually, when we consider the usage of infrastructures, its efficiency, reliability and dependability are key parts. In general, big sensing data are stored in the cloud. However, in disaster situations, it might not be able to access to the cloud from disaster areas. Thus, it is important to consider the efficiency, reliability and dependability for not only the cloud side but also the sensing edges. In order to design and develop mission-critical services, we definitely need to consider failures of communication lines.

A kind of autonomous recovery from such failures should be equipped. Also, data security and privacy are both important. Dissemination of incorrect information and false rumor might make the society confusion. Given the large volume of data and the real-time constraints for performing the analysis for prediction as well as the analysis for appropriate response and recovery operations, it is imperative to ensure that the computations can be done in real-time. A variety of methods may have to be integrated together to achieve this, including the use of parallel processing based on cloud computing and local resources, reduction of large data sets into equivalent correct rules, distributed pipeline processing, etc. It is also important to ensure that the data acquisition and analysis procedures are highly dependable in spite of the failures of various processing and communication units.

Given the distributed nature of such computing, it can be difficult to identify which units have failed. Hence, the processing and communication infrastructure may need to be augmented with dependable on-line system health monitoring capabilities to enable the rapid identification of faulty components and the activation of redundant standby units to ensure correct and timely completion of the big data analysis under emergency situations. Assessment of the big data analysis algorithms is needed to determine the confidence in the correctness of the results of the analysis, including predictions and recommendations for optimal response and recovery actions. Simulation and emulation platforms can be used for evaluating and certifying new algorithms and procedures. Different algorithms can be evaluated and compared by applying them to a suite of benchmark scenarios and test-beds for which the correct results are known. It is important to quantify the confidence in the accuracy of the results of big data analysis since failure to predict a disaster in a timely way can be harmful to society. Similarly, false alarms should also be avoided since these can reduce the likelihood that the public will react appropriately to a real disaster.

2.2 Volume and Velocity

In many cases, data generated by sensors need to be processed in real-time for immediate action (e.g., Twitter and Facebook). However, the development and validation of models using real-time data is a challenging task. Archive of datasets collected during historical events that can be shared among researchers would be helpful in enabling quantitative analysis between different models – i.e., the use of a common dataset for evaluations and validations of models would help researchers in developing models with improved quality. Open access datasets collected during real events are known to be very useful to test and validate new ideas: for example, the 1998 World Cup web access trace [6] has been used by many research to advance web server/services technology.

Research on cloud-based approach can potentially lead to a viable solution to accommodate the high volume of data and the multiple formats of data generated by different sensors of various types, While cloud systems have the necessary components, namely compute, storage, and network, to the development of a repository of disaster-related datasets, models, and applications, research is needed to integrate components and design interfaces that make it easy for domain scientists – which are not necessarily cloud computing experts. The big data gathered through sensor and social networks will become useful once they are turned into actionable information that helps decision-making. There is multi-scale timeliness to decision making during disaster response. Some information should be available on the order of seconds, minutes, and other can be a matter of hours. Faster is better, but, high error rate could aggravate the situation. For example, inaccurate information could result in the distribution of rescuers and supplies to wrong places, wasting limited resources. False information could also misguide the public, increasing their stress level.

2. 3 Variety

Integration of Many Heterogeneous Data Sources The integration of many heterogeneous data sources and software tools when applying big data to disaster management is a significant challenge. At a limited scale, this heterogeneity is a challenge that has risen in big data scientific research, for example, in the construction of Global Climate Model (GCM) for global weather modeling. To achieve global weather prediction, a GCM needs to integrate a variety of atmospheric and ocean models. Given the relatively small number of component models (less than a hundred), it has been feasible to connect them manually. In contrast, the number and variety of data sources in disaster management, as well as their rate of change, far exceeds what is feasible by manual integration.

Consequently, research on automated development and maintenance of data integration tools is necessary and its success very important. Big data analytics for disaster management and response requires a large variety of heterogeneous data sets that are related with each other and show different aspects of the changes caused by a disaster. We need the integration of such heterogeneous datasets for big data analytics. We need to handle many kinds of sensors outputting different types of data ranging from time series data to semi-structured data and textual data. These data inherently include noise and misinformation. We need to improve the trust and reliability of these data despite some noise and misinformation in them. For example, by combining information from multiple, potentially unreliable, but independent sources, we may statistically improve the trust and reliability. We also need to take into account that there may be some dependency within and among many resources. Re-tweets, for example, are not mutually independent.

2.4 Metadata Management Issues

For an appropriate interpretation of heterogeneous big data, detailed metadata is required. Some of the reports contain some metadata, but many more details (e.g., about the specific sensor used in data collection) are needed for research purposes. The collection of metadata and data provenance is a significant challenge when the data are collected under duress and stressful situations.

Furthermore, the sensors are operated by a large number of different government agencies for different purposes. At the national level in Japan, the agencies include: Ministry of Internal Affairs and Communications (Fire and Disaster Management Agency); Ministry of Agriculture, Forestry and Fisheries (e.g., radioactivity in fish); Ministry of Land, Infrastructure, Transport and Tourism (Land and Water Bureau); Ministry of Environment; Ministry of Health, Labor and Welfare (Pharmaceutical and Food Safety Bureau). Many more sensors are operated by prefectures, universities, and other agencies. Consequently, the main challenge has shifted from sensor insertion for a relatively small amount of data collection to the management of large but varied data from many sensors.

2. 5 How Emergency Managers Can Benefit from Big Data

During a disaster, life-saving decisions are often made based on the most current information of a situation and past experiences in similar circumstances. While that's a tried-and-true approach, the availability of complex, computer-generated data streams is changing the ball game for some emergency managers. Large volumes of data sets — commonly referred to as big data — derived from sophisticated sensors and social media feeds are increasingly being used by government agencies to improve citizen services through visualization and GIS mapping. In addition, big data is enabling responders to react to disasters more efficiently.

Volunteers at Splunk, an operational intelligence software provider, are involved in a project that culls data from Twitter feeds. By analyzing keywords along with time and place information, a pattern of activity in a particular area can be unearthed. The idea was used during Superstorm Sandy. FEMA created an innovation team composed of public agencies and private companies. One of the participants was Geeks Without Bounds, a nonprofit humanitarian project accelerator, which partnered with Splunk's charity arm, Splunk4Good, to apply the social media analysis. Team members working on the project looked at hashtags and words in Twitter feeds as well as Instagram photos related to Sandy, evacuation rates in specific areas and other keywords about resources, such as power, food, fuel and water. Using that data, the team plotted out locations where supplies might be most needed and got a finger on the pulse of a community's sentiment about available resources.

"You can imagine the ways it can be used in real time for response during an emergency," said Stephanie Davidson, director of federal civilian sales for Splunk. "It's really helpful for where to allocate those resources to the places that need them most."Government agencies have been using social media data for sentiment analysis and public relations for a while. But according to Art Botterell — associate director of the Disaster Management Initiative at Carnegie Mellon University, Silicon Valley — practical use by emergency management agencies for response, recovery and preparation activities is fairly new. Botterell called current efforts of emergency managers using social media a period of rich experimentation, where decision-makers must determine whether big data derived from Twitter and Facebook should be further incorporated into practical emergency situations, or used simply as a communication tool. “This is an area that has been technology- and concept-driven, which is how most innovation happens, but now we’re getting to the point where it all falls under the big data tent [and] how do we know what is more useful and less useful,” Botterell said. “This is a conversation that I haven’t heard emergency managers having.”

3. Big Data in Disaster Management

Effective Big Data Disaster Management

Below are some key points to consider when setting up big data disaster recovery and management:

1. Regulatory Requirement

Your big data disaster management plan needs to comply government and relevant regulatory body mandates. Some companies like knowledge and information aggregators, who look at long term trends, may keep even decades-old data, while this is not necessary for other type of companies. Blend in your critical time period for retaining data, with regulatory compliance essentials. This detail will help you to correctly determine how many years back you need to go, when preserving data.

2. Recovery Point

Big data is a culmination of various well thought out processes. First, the use of a web data extractor collects relevant and targeted data from diverse sources. This is then stored in a data warehouse in a planned manner. From here, it is passed through an ETL engine (extracted, transformed, and loaded) to be used by BI and big data analytics tools for uncovering insights. Now, coming back to big data disaster recovery, you need to be sure what will be your recovery point in case of an outage or data issue. Will it be the initial raw format that data comes in from various sources? Or will it be a more refined form of data that has passed through the

ETL process? A larger proportion of corporate entities involved in big data disaster management will choose the second option. It is worth noting that this is just like transactional data – where the point of recovery will be nearest to when the stoppage or trouble happened in the transaction. However with big data, it is also necessary to determine the 'form' in which you need the data to be recovered.

3. Speed of Data Recovery

In order to derive its full potential, management rightly needs big data analytics to be carried out in near real time. This calls for quick fire recovery in case of any trouble. IT heads might look at cloud storage options to enable this. They can also look to bolster their on-site storage choices. This can be done either with a slower media such as tape drive that has time consuming recovery, or go for continuous replication on in-memory storage on more than one data server.

4. Priority vs. Non Priority

Big data is humongous. You need to be absolutely clear on what data takes priority for recovery, in case of an unforeseen disaster. It's unnecessary and expensive to try and recover ALL the data at once. Mission critical, time critical, or rapidly changing data needs to be on top of your priority list, followed by other data constituents that change or get updated less rapidly. Classifying various clusters into these two types of data on priority needs to be a consensual affair. This way the top management, operations heads, IT admins, and other stakeholders can mutually decide on what should be recovered first and what can wait for a bit longer.

5. Enforce Data Governance

The classic quip of 'Prevention is better than cure' becomes very applicable for big data disaster recovery. While we may have strong process oriented approach for related activities such as data crawling, we lag behind when it comes to implementing strong protocols for big data. Having this in place will help with data disaster management to a great extent. Important components worth considering are assessing the data provenance (source and metadata about the data) and then deciding on how to use the data in your analytics. This is not something that you can simply plug and play. It has to be ingrained into the way a web data extractor works. This helps make certain that you consider this crucial pointer from the very outset of data collection and later on, its analysis.

6. Practice Makes It Perfect

All the initial brainstorming and devising of big data disaster management plans will bear fruit only with its successful long term implementation across the organization. All stakeholders (both, internal and external) and IT executives need to be taken through the plan and told of their roles and responsibilities in context of the larger picture. It needs several rounds of testing and coordination with end users (individual as well as department level) and external vendors.

The involvement and adoption will be hesitant from them initially, but don't lose hope, because this is the foundation to your disaster management initiative success. You will see that they will eventually support this planning and preparation process when they witness the immense value provided by a properly planned big data disaster management process. These pointers will help you ensure that your big data continues to denote integrity, accuracy, and relevance to help management with insight-based decision making. Do write in to us and let us know what consideration have you factored in for ensuring the safety of your hard earned big data insights.

4. Motivation and Overview of Big Data Analytics

Although many issues for ECNs have been studied, some research problems are still open. [4] It is hard to understand global situations, e.g., a group of people with similar movement patterns. Such kinds of information are important in ECNs when considering the limited communication resources and dynamic changing environments. This information can possibly be grasped by using big data analytics. Efficient deployment and adjustment of different types of

ECNs requires understanding the global situations in disaster areas. ECNs are different from general cellular communication systems:

1) Disaster context awareness,

2) Infrastructure free systems,

3) Distributed style,

4) Limited communication resources,

5) Requirement of optimal deployments and adjustments based on situations as they occur, etc.

Possible data analysis techniques, e.g., stochastic modeling and data mining, were briefly reviewed for wireless communication. However, they neither address the new challenges from the unique characteristics of ECN from both content and spatial points of view. Also, case studies are not sufficient and detailed enough. Therefore, in this paper, we perform a systematic survey and an attempt study on big data analytics for ECNs, through discussing analysis methods and their applications from both content and spatial points of view, and several detailed case studies. From a content point of view, an analysis focuses on the content of the data, e.g., personal opinions, topics, and sentiments. From the spatial point of view, both content and its spatial information are used together in an analysis to explore the distribution patterns of the data in space.

Analyzing such data can prompt a better understanding of the situations when disasters take place. However, both kinds of big data analytics follow similar procedures, shown below:

Collection

The first step is to universally collect the data. The analysis process cannot be conducted without data, and the appropriate data are especially important.

Storage

The second step is to maintain large-scale data storage servers. The size is not a problem with the current mature technology, but a real challenge will be efficient processing of the distributed data which also covers a wide spectrum of issues such as scheduling problems on single computers or open clusters, communication guarantee or failure recovery, and data transmission between vertices.

Management

The third step is to correctly categorize the data. Since big data are often from heterogeneous resources, applying flexible data management offers a more reliable and instant storage plan for heterogeneous information at different levels, compared to general data management solutions. It also mass data operations suited for further analytics and statistics.

Processing

The fourth step is to process and analyze the data in an efficient way. To achieve instant large scale data processing, it is necessary to use distributed computing resources.

Prediction

The last step is to make a prediction. which makes a meaningful connection among separate datasets for a specific purpose, based on the results obtained from the raw data. This step also requires the participation of domain experts and users, who are expected to observe and explain the meanings of the data and determine the solutions (as well as an action plan).

5. Big Data: A Natural Solution for Disaster Relief

With big data as common in science as it is everywhere else, could we have used better tools to see this coming? What's the role of big data in natural disasters today?

The answer is a work in progress. NASA, for one, admits to currently having a big data problem. “(D)ata is continually streaming from spacecraft on Earth and in space, faster than we can store, manage, and interpret it,” writes NASA Project Manager Nick Skytland. “In our current missions, data is transferred with radio frequency, which is relatively slow. In the future, NASA will employ technology such as optical (laser) communication to increase the download and mean a 1000x increase in the volume of data. This is much more then we can handle today and this is what we are starting to prepare for now. We are planning missions today that will easily stream more then 24TB’s a day. That’s roughly 2.4 times the entire Library of Congress – EVERY DAY. For one mission.”

NASA still needs to catch up with its data load. Other government agencies are looking for ways to collaborate more effectively. For example, the Department of Defence has secret satellites located around the world for reconnaissance. Those satellites also happen to have the capability to detect large and small meteors. The DoD, however, is nervous about sharing any information that it deems classified, so efforts are still underway to find a way to incorporate that data into the bigger scientific schema.

6. Real-Time Disaster Maps

Terrestrial challenges, on the other hand, are currently more amenable to big data. One of big data's true strengths lies in crisis mapping, the process of using visualizations, footage, analysis and apps to get an overview of a disaster as it evolves. Google's Super storm Sandy Crisis Map tracked the course of last winter's storm, with video footage, evacuation routes and emergency aid centres. The UN commissioned the Digital Humanitarian Network to track the real-time effects of Typhoon Pablo in the Philippines. Among other efforts, social data was analyzed to provide a detailed, real-time map of displaced people, fatalities, crop damage, broken bridges and more.

7. Using Big Data In A Crisis: Nepal Earthquake

There were hundreds of emergency services, charities, disaster relief agencies and volunteers have done their best to help people affected by the terrible Nepalese earthquake which struck during the weekend. And Big Data is playing its part, too – with crowd sourced, data-driven efforts to connect people outside the country with their missing loved ones, and assist in getting aid to where it is needed.

Much of the work on developing Big Data systems to help with disaster relief began in the wake of the 2010 Haiti earthquake and the 2011 Tohuku, Japan earthquake and tsunami. Japan and the US instigated a joint research program to find workable methods of using data to ease the toll of natural disasters which kill thousands each year, and cost the global economy billions. Last year, the US National Science Foundation and Japanese Science and Technology Agency offered $2 million in funding to groups working on data-driven solutions to disaster management problems. At the other end of the scale, crowd sourced, data-led initiatives have also started off at a grassroots level, with community members coming together to collate data to assist others. This happened following Hurricane Sandy in the US, when high school students collaborated to create an online map of the New York and New Jersey area showing where gas was available. Following Typhoon Haiyan in the Philippines, the international Red Cross collaborated with volunteers around the world to map the effects on the region and its people.

The four key elements of disaster management are prevention, preparation, response and recovery. Big Data has potential to help with all of them.

While not much can be done to prevent natural disasters, sophisticated Big Data systems such as those developed by Palantirare being used to crack down on man-made disasters such as those caused by terrorism. But when it comes to "acts of God", of course the focus will be on preparation, response and recovery.

8. Adoption Issues for Big Data Approaches in Disaster Management

One of the key questions in disaster management research is the acceptance and adoption of technologically advanced solutions by the public. To illustrate improved emergency response, let us consider the idea of "call in emergency" list, which has been implemented by many institutions. For example, large institutions such as University of Florida and Georgia Tech offer an opt-in emergency notification service to send critical information to employees and students when emergencies arise. The notifications modes supported currently include phone calls, emails, and SMS to smart phones.

9. Practical Questions on Technology Adoption

Another emergency response scenario being suggested (and debated) is the use of smartphone location information to notify every phone in a danger zone during an emergency. (For the following discussion we will assume that the legal rules about opt-in and opt-out can be satisfied.) Although the idea of emergency notification is a relatively simple concept, there are many opportunities that arise beyond a hardwired message. Let us consider its application to the 3-11 scenario, where the tsunami arrived the Tohoku coast a few minutes after the earthquake. A quick notification (e.g., SMS) could be sent immediately to the coastal areas after an earthquake has been confirmed. As more information becomes available, e.g., the epicenter of earthquake is determined, phones are classified by their distance to the epicenter and their risks assessed accordingly (e.g., their distance to the seashore. Additional emergency response information is then sent to phones according to the risk level due to their location. This additional information can be customized for each location as well as other available information such as whether the phone has been used or moved since the start of earthquake. For phones with active users, appropriate escape routes could be sent to each phone, using facilities such as Google Crisis Response. Each refinement step requires significant knowledge of the environment (provided by big data sources) plus an accurate knowledge of the disaster for accurate risk assessment and appropriate response. To increase the adoption of technologies developed for disaster prevention, preparation, response, and recovery, it is essential to consider the integration of the new infrastructure with tools currently in use and offer familiarization of new tools before a disaster takes place. For example, geographical replication is a known solution to recover IT from a devastating disaster.

However, these incur costs that prevent many small businesses from implementing it. Given that electronic businesses are increasingly adopting cloud computing and virtualization, [7] proposed the use of migration to mitigate the impact of disasters to an IT infrastructure when a disaster is predictable, while [8] increases the possibility of adoption for those taking advantage of advanced I/O technologies in virtualized environments. Studies that consider the barriers for adoption such as incompatibilities between solutions, cost/benefit of a solution, cultural differences, and amount of training are expected to accelerate adoption by providing accurate assessments. It is important to have opportunities for public training. However, there are so many people who are not familiar with ICT devices. Thus, we should develop ICT devices with simple and understandable user interfaces so that many people can recognize emergency alerts.

To increase public awareness in disaster prevention, preparation, response and recovery, it is critical to create effective training environments and programs that help prepare the public for future disasters. This can be achieved through appropriate education programs, such as ones that teach the public about potential disasters and how they progress, how should one response in such disaster situations, what tools, services and help are available, how one can effectively assist others during actual disasters and post-disasters. In addition, conducting regular testing of the services and public responses under simulated disaster settings is important not only to increase public knowledge and awareness but also to evaluate the services and infrastructure. Collaborations between various government and industry stakeholders and scientists of various fields (e.g., engineering, psychology, healthcare) to construct such training environments and programs are needed to effectively educate the public.

10. Conclusion

Disaster Management is an important global problem. Disasters affect every country on Earth and effective disaster management is a global challenge. This is particularly the case of large-scale disasters that affect many countries (e.g., the 2004 Indian Ocean earthquake and tsunami) and multi-hazards such as the Tohoku Earthquake and landslides. Big Data is a great global opportunity for disaster management. Big data has already demonstrated its usefulness for both dedicated sensor networks (e.g., earthquake detection during the Tohoku Earthquake) and multi-purpose sensor networks (e.g., social media such as Twitter). However, significant research challenges remain, particularly in the areas of Variety of data sources and Veracity of data content. We call this the Big Noise in Big Data challenge. From the disaster management view, we need the technology push from big data researchers to tackle the challenges mentioned above (e.g., Big Noise) so big data tools can effectively address disaster management issues. From the big data view, we need the application pull of disaster management researchers to apply big data techniques and tools to solve real world problems.

==

References

[1] S. Kailser, F. Armour, J. A. Espinosa and W. Money, "Big Data: Issues and Challenges Moving Forward," 46th Int.Conf. System Sciences, pp. 995.

[2] S. Kailser, F. Armour, J. A. Espinosa and W. Money, "Big Data: Issues and Challenges Moving Forward," 46th Int. Conf. System Sciences, pp. 996-997.

[3] S. Kailser, F. Armour, J. A. Espinosa and W. Money, "Big Data: Issues and Challenges Moving Forward," 46th Int. Conf. System Sciences, pp. 996-997

[4]. Junbo Wang, Yilang Wu, Neil Yen, Song Guo , and Zixue Cheng." Big Data Analytics for Emergency Communication Networks: A Survey" Journal Of Latex Class Files, Vol. 13, No. 9, September 2014

[5]. C. Pu and M. Kitsuregawa. "Big Data and Disaster Management A Report from the JST/NSF Joint Workshop"Technical Report No. GIT-CERCS-13-09; Georgia Institute of Technology.

[6]. M. Arlitt AND T. Jin, "Workload characterization of the 1998 World Cup Web site," Technical report, Hewlett-Packard Laboratories, September 1999.

[7]. Tsugawa, M.; Figueiredo, R.; Fortes, J.; Hirofuchi, T.; Nakada, H.; and Takano, R., "On the use of vir-tualization technologies to support uninterrupted IT services: A case study with lessons learned from the Great East Japan Earthquake," in 2012 IEEE International Conference on Communications (ICC), 2012, pp. 6324-6328.

[8] Takano, R.; Nakada, H.; Hirofuchi, T.; Tanaka, Y.; Kudoh, T., "Cooperative VM migration for a virtual-ized HPC cluster with VMM-bypass I/O devices," E-Science (e-Science), 2012 IEEE 8th International Conference on , vol., no., pp.1,8, 8-12 Oct. 2012. doi: 10.1109/eScience.2012.6404487.

==

Jebas Sinthiya. I., PG Scholar
Computer Science and Engineering
Anna University Regional Campus
Madurai 625 019
Tamilnadu
India

E. Shanmugapriya
Assistant Professor
Department of CSE
Anna University Regional Campus
Madurai 625 019
Tamilnadu
India
vijaylaya2000@gmail.com