PAPER • OPEN ACCESS

Disaster Management in Smart Cities using IoT and Big Data

View the article online for updates and enhancements.

Disaster Management in Smart Cities using IoT and Big Data

Akil Ramesh, S Rajkumar[*]**, Jenila Livingston L M**

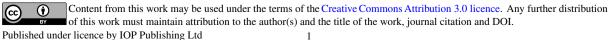School of Computer Science and Engineering (SCOPE), VIT, Chennai, India - 600127

Email: *rajkumar.srinivasan@vit.ac.in

Abstract. Disasters are the most vulnerable factors which are affecting the human life and environment. Smart cities are mainly affected by disasters. This shows the requirement of an efficient disaster management system in smart cities. This paper deals with disaster management technique used in smart cities to detect disasters like building fire, pollution in atmosphere, route blockage using Big Data Analysis and Internet of Things (IoT). A combination of Hadoop environment is used in disaster reliant smart cities. IoT is used as a front end for collecting real time data sets and Big data analysis for data pre-processing and aggregation. This work evaluates the threshold capacity of each data which is collected for these disasters using Z-score normalization and if collected parameter exceed above the threshold value, then alert message is sent to public service sector such as fire station, police station and hospital. Different challenges faced by this evaluation are discussed in this manuscript.

Keywords: IoT, Big data Analysis, Disaster, Z-score Normalization, Fog computing, Cloud computing, Ad-hoc network, Public services.

1. Introduction

Intensity of natural and man-made disasters is increasing day by day. Smart cities are mostly affected because of rigid construction of building, high population and high traffic. So, the effect of disasters in cities will be worse comparing to village. Also, the casualties and damages in cities will be very high Effective disaster management technique should be introduced in smart cities. In [2], mainly use two criteria's Big Data analysis (BDA) and Internet of Things (IoT). IoT Devices are sensors used for collecting real time data sets from the environment (structured data) and twitters are used to get non structured data. IoT allow communicate within the devices via internet, as it will increase the data collection and data transferring procedure efficiency. But the problem with IoT devices in disaster management technique is they collect the data or parameters daily, therefore the data storage and management tend to be very challenging. In order to avoid this situation, we use Big Data analysis as the back end for processing data for the IoT data. For BDA important task is to find hidden values from data sets having big size of various data types. Big Data analysis is a modern efficient way for processing data. In this paper, we use Hadoop environment and spark platform to aggregate and segregate data according to the similarities in their behaviour in thee storage place by processing the obtained data.
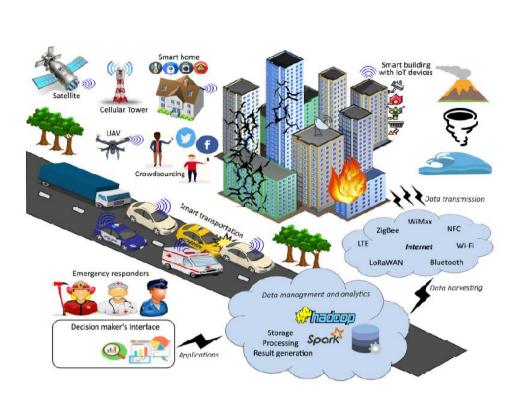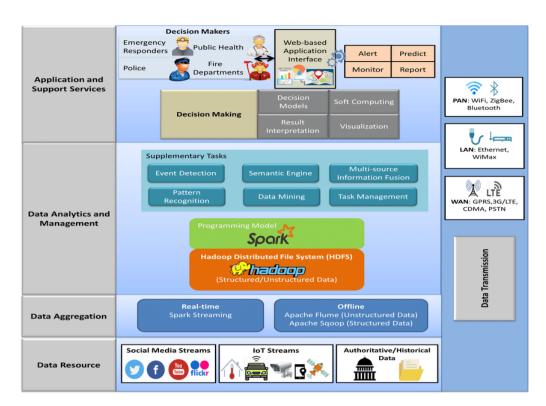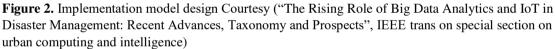
Figure 1. Disaster management environment using BDA and IoT Courtesy ("Towards Disaster Resilient Smart Cities: Can Internet of Things and Big Data Analytics Be the Game Changers?", IEEE trans. on special section on urban computing and intelligence.)

Figure 1 shows how data sets are taken from various objects and how Hadoop and spark system involved in data processing and various IoT devices which are used to collect data sets from real time system and unstructured data from social media and their interconnections. Disaster management technique in [1] deals with smart cities using IoT and BDA. This paper focuses on how an IoT device and Big Data analysis can be grouped to create an efficient disaster management system. The proposed work [1], mainly focused on pollution control, building fire control and traffic management system. The implementation model consists of five layers such as data resource layer, data transmission layer, data aggregation layer, data analytics and management layer and application and support services. In data resource layer collection of structured and unstructured data will be collected from IoT devices and social media mainly Twitter data. In data aggregation layer real-time data are aggregated using spark streaming platform and unstructured data are aggregated using apache flume and structured data using apache scoop. In data analytics and management layer it will segregate the data using Hadoop and spark environment. In application and support service layer, it will compare given mean value of parameters with threshold and if it exceeds, it will send an alert message to public service systems. Data transmission layer is used to inter communicate with each of the above layers using LAN, WAN and PAN. Each layer has its own functionality which is dependent to each other layer. This makes the system more efficient disaster management system of modern times.

Figure 2. Implementation model design Courtesy ("The Rising Role of Big Data Analytics and IoT in Disaster Management: Recent Advances, Taxonomy and Prospects", IEEE trans on special section on urban computing and intelligence)

To improve the data resolution efficiency, we use spatial clustering technique. Spatial clustering technique is basically worked in fog computing environment which is an intermediate layer below cloud. Usually network connection loss will occur, in order to overcome that there is Emergency Communication Network (ECN) is used. Also, it has some mobile base station called Movable Base Stations (MBSs) which come into play when disaster alert occurs. Spatial clustering analysis is a technique used BDA to aggregate and cluster data efficiently based on their similarities in the feature.

2. Related Work

The working system in [3] mainly deals with how to prevent delay in transferring data from our data storage system after processing to the public service system. At the time of disaster occurrence, we will not get stable network signal to transfer data or report emergency. Rather than selecting cloud storage technique there is an intermediate layer called fog computing. Fog computing have an effective emergency communication network comparing to other technique. It has a movable base station which has node to node communication. This base station comes into action whenever it is necessary. They are mainly used for disaster area adapted delay minimisation. This is a type mobile computing technique where base station act as a portable platform. In previous techniques there will be some delay due to signal problem or accessibility issues.
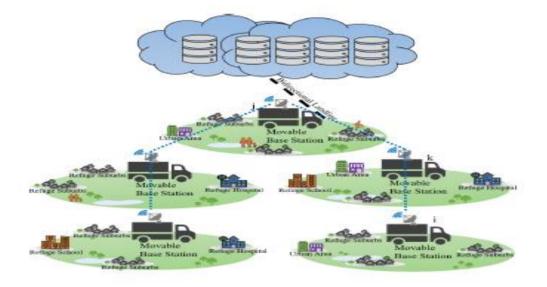
Figure 3. System model Courtesy ("Maximum Data-Resolution Efficiency for Fog-Computing Supported Spatial Big Data Processing in Disaster Scenarios", IEEE trans. on fog and cloud computing)

Here, single node which has computation and communication is evaluated. With this node smartphones communicates and upload large amount of data based on user's usage and it will move to fog layer and does BDA and data are clustered and conclusion is transferred quickly. There is no wired communication happening, data are transferred wirelessly. This is a type of efficient data transferring technique used to minimise the delay due to signal problem. Fog computing enables a stable communication between different node, which increase the efficiency of disaster management technique communications.

Spatial Big Data Analysis is very important in disaster management scenarios to understand distribution patterns of the situation. Formulating clusters from cloud computing always take time, so we are using spatial data clustering in fog computing. Spatial clustering is a technique which is used to cluster data with similarities in their characteristics and behaviour. They also categorise dissimilar data in another cluster, so we can evaluate and check their credibility and we can remove that data accordingly. They will efficiently cluster useful data as a cluster so we can minimise the time required for data processing, which will increase the efficiency of the management system.

The proposed clustering system in [4] has 2 functions mainly Spatial Data Aggregation Function (SDAF) and Spatial Data Clustering Functions (SDCF). First, it will divide the data into m X n cells then it will do spatial aggregation of data based on the similarities in their properties. After aggregation the data are met with spatial clustering function with will group the data into matrices according to the needs. Density Based Spatial Clustering of Application with Noise (DBSCAN) is a type of clustering method in which data are clustered based on the distance between the radius or distance between them based on their similarities in parameters. If the size of data is very large, we use another technique called grid-based clustering technique. In this technique if the size of data is very large then they are clustered into grid like mesh structure. This grid-based technique scans all the possible mesh classification in the grid and finds the correct related data in the grid which is useful for evaluation. In this system, there are 3 types of data namely, raw data is type of data which we collect from real time and from other sources. Second type is point data which is grouped after analysing and abstracted as predefined labels and final is page data which is created by aggregation of point data.

In [6] deals with IoT and fog computing based evacuation technique used at the time of disaster. Previous evacuation techniques are mainly based on non-humanitarian aspects, but Crowd Lives Oriented Track and Help Optimization system (CLOTHO) mainly consider psychological and humanitarian aspects of human beings and plan the evacuation accordingly to individuals. The main objective of using IoT devices and social media data of respective person is to identify the behaviour and psychological behaviour of each individual. CLOTHO mainly deals with Artificial Potential Fields (APF). APF is conceptualized as an IoT services and can determine the direction of evacuation to the potential field. APF for every individual will be processed. When a disaster occur an artificial rescue system will be initiated on the mobile phones from the cloud services according to their potential field attributes. The main purpose of mobile computing is from the loud evacuation technique will be delivered to each one's mobile phones and it also shows the nearby shelters available, easy routes for rescue and nearby medical camp place considering humanitarian aspect.

CLOTHO mainly has four modules in the architecture. First module is data acquisition module in which location and personal information of evacuation is collect through mobile phones. Second module is the network transmission module mainly responsible for two-way communication between the mobile devices and cloud services. In [5] it mainly deals with flood management system using IoT and BDA. This paper ensures how to implement a flood management system efficiently. For this efficiency they ensure the correlated factors also in a well-planned manner. Here for placing the IoT devices hexagonal classification of devices is proposed such that a high prone location is detected and devices are placed in such a manner that x axis goes from highest area to lowest prone area where as y axis catches the width. Each device is placed in the distance of half the radius of catchment, so the evaluation or data collection will be more accurate. It also use energy saving algorithm for IoT devices in such a way that it will detect when threshold value is greater than the threshold to switch on the devices, it will alert else it will remain idle. Single value decomposition (SVD), Holt-Winter's forecasting and k-mean for evaluating collected data. Using single value decomposition the collected data is decomposed to a single mean value. This decomposition helps for easy evaluation of data. K-mean clustering is used to predict current status of the flood. After clustering if new data arrives, it not only defines its cluster but also upgrades itself by minimising cluster distance. Holt-Winter's forecasting technique is one of the best forecasting techniques because every year there will be variation in total precipitation therefore it is very difficult to maintain same database. Therefore, in this work it will compare the trend, seasonal precipitation, period after which the value started repeating, time period, current situation. Mean value of these situations are taken and the forecasting is evaluated.

From [2] it is clearly evident about the advantages and disadvantages of using IoT and BDA based disaster management system. By using IoT devices, we can collect real time data sets from the atmosphere which helps for frequent data collection and faster mode of data collection for analysis. IoT devices can be easily installed without affecting any human life and daily routines. IoT devices can be easily connected to internet sources to cloud and other storage platform for analysing. Also it will have good interoperability between devices established throughout the surroundings. IoT is a device which is also known as multiple data source devices. BDA is a latest kind of analysing technique to analyse large volume of data frequently. BDA allows high speed data analysing, Real time query generation and high speed result generation. Here, we conclude for faster data collection hence IoT devices are essential for analysing large volume of data. In [7], a model was developed to deals with post-disaster management system using IoT devices. This paper mainly specifies the importance of IOT devices for post management and how it is helpful in real time. After a disaster happens, there will not be any proper mobile network for communication. To solve this issue IoT

devices are installed in the environment to connect with the system and detect the disaster information.
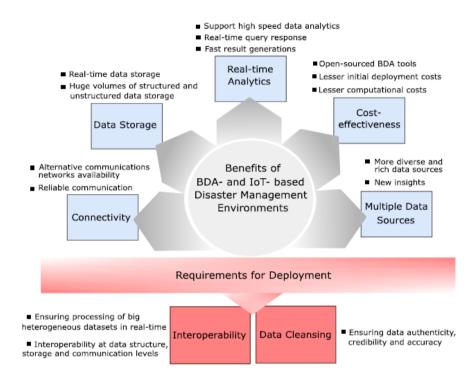

Figure 4. Advantages and disadvantages Courtesy ("The Rising Role of Big Data Analytics and IoT in Disaster Management: Recent Advances, Taxonomy and Prospects" , IEEE trans. on urban computing and intelligence)

If there is a connection loss it will use Device to Device (DtD) Ad-hoc network which act as an agent to connect to the server and transfer data for analysing .These relay agent is used to establish the connection between the devices. This connection will mainly depend upon the factors like computational power, residual battery, channel quality index, bandwidth availability etc. The paper [8] deals with various big data tools and why big data is used. For disaster scenarios we need large data and normal storage system won't be sufficient to deal with. To resolve all these issues we use big data analysis technique which will help to store large amount of data and also analyse these large quantity of data easily. Big data paradigm consists of data management tools and techniques for storage, processing and security of data [9] deals with the challenges faced by smart sensing devices like IoT devices to collect data from the location. The main challenge here is the lack of connectivity at the time of the disaster. If disaster happens, then interconnectivity between IoT devices will occur, therefore the IoT response will damage at that time. In this paper this mode of data collection is named as smart sensing technique.

Figure 5. People centric data collection Courtesy ("Smart disaster detection and responses system for smart cities", IEEE trans. on computers and communications)

After data collection challenges, they are raising the challenges to data processing. This data processing should be done with proper algorithms. The volume data is too high so data processing should be normalised without diminishing the significance by using different subjects like machine learning, big data analysis, data mining, data indexing etc. This system is known ad smart processing. After smart processing it should be made notified to public. This phase should be done more accurately in less amount of time. At the time of disaster there will be lack of network to communicate, As a solution for this problem, use of ICN (Information Centric Networking) technique has been raised. ICN uses content name rather than server location as communication entities. Big Data Analysis and data mining is the technique which is used to analyse data and to extract hidden information.
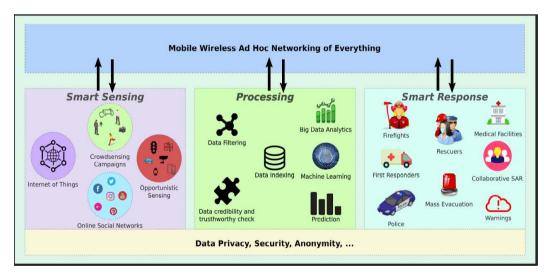

Figure 6. Smart system Courtesy ("Smart disaster detection and responses system for smart cities", IEEE trans. on Big Data Analysis)

3. Proposed Work

The proposed system of disaster management in smart cities using IoT and big data analysis provides a system in which data collection using IoT devices from the real time environment. There are two different types of data which can be collected such as structured data that is collected from various devices and unstructured data which is collected from social media and human sources can be used in this system. After the data is collected it is provided with Big Data analysis using Hadoop framework. In this paper, mainly evaluation of three disaster condition is examined in a smart city such as pollution in a city due to high amount of certain gases like carbon dioxide, carbon monoxide, sulphur dioxide, building fire due to increased temperature, traffic which can be caused due to disaster. This disaster environment uses Hadoop framework in Big Data analyzation to analyse the data collected from various sources because these data will be in large volume so it is very difficult to analyse data in other environment and to store these data will also become a challenge. This large volume of data can't be evaluated or analysed as it is because it is very difficult to analyse these huge volume of data and thee will be many noises and variation in these data which leads to wrong results. Therefore to overcome this problem we are using map-reduce technique with Z-Score normalization technique. The drawback of normal mean normalization does not contain any attribute to calculate the deviation in the data thus causing challenge for accurate results. Hence, the normal mean normalization is not used in this work as it takes the average of the data and plot. But, it will not remove the noises or evaluate the deviation in the data. Hence, to overcome this challenge we are introducing Z-score normalization technique for the normalization of data. Z-score normalization will calculate the value by dividing the standard deviation of the data set.

3.1 Z-Score normalization

Z-Score Normalization is represented using the following relation:

$$z = (x - \mu) / \sigma$$

Where, Z is the normalized value, X is the value of individual data in the data set, μ is the mean of the distribution and σ is the standard deviation of the dataset. Z-score normalization will provide the data in the precise manner in the range of 0~1 where 0.5 is the threshold value to determine disaster and normal condition from the data set.

3.2 Pollution Alert

In this module, we determine the pollution level of the city by calculating different gas level in the atmosphere. The data is normalized using z-score normalization. The data will be normalized in the range of 0~1 using this normalization from which 0.5 is the threshold value for evaluation. If the data range is above 0.5, then alert message will be sent to public service sector.

Algorithm-1: Pollution alert

Begin

Input: gases data

Output: pollution alert/no pollution

1. for each gases loop

2.   if gases value > 0.5 (threshold value

     from z-score normalization)

3.   pollution alert

4.   else

5.   Go to (Step1)

6.   End if

7. End loop

3.3 Fire Alert
This module determines the probability of catching fire to a building by evaluating the temperature level of the building. Collected data is normalized using z-score normalization in the range of 0~1 and the threshold value is calculated as 0.5 in z score normalization. All the value above 0.5 will notify the public service a fire alert and below 0.5 as normal.

<u>*Algorithm2: Fire alert*</u>

Begin

Input: Temperature data

Output: fire alert/normal

1. for each temperature value loop

2.   if temperature value > 0.5 (threshold value from z-score normalization)

3.   fire alert

4.   else

5.   Go to (step1)

6.   End if

7. End loop

3.4 Blockage Alert
This module determines the traffic caused due to any disaster. It will analyse the number of vehicle in the location and provides traffic alert if the value is greater than the threshold value. Traffic data is normalized similarly using z-score normalization.

<u>Algorithm 3: *Blockage Alert*</u>

Begin

Input: blockage data

Output: blockage alert/normal

1. for each blockage value loop

2.   if blockage value > 0.5 (threshold value from z-score normalization)

3.   blockage alert

4.   else

5.   Go to (next reading)

6.   End if

7. End loop

4. Results

In this paper, the disaster management system is developed by using IoT and BDA with Z-Score normalization. The results shows this proposed approach is more accurate and efficient comparing with a system using normal mean normalization because it doesn't involves any deviation of data. In the earlier systems there can be large chances of results deviating from the original data.

In the existing system, mainly they followed normal mean normalization which computes the mean value of the data set as threshold value. But, the data from IoT will be very large in quantity and even there will be some noise data in the collected data set. It is not easy to understand the noise data from this large volume of data. This will lead to deviation of data, which will cause inaccurate analysis. In the existing system, the threshold value changes based on readings, so early prevention is impossible. Hence, in the proposed system we will use Z-Score normalization where there is a fixed threshold after calculating the deviation in the data sets. Therefore, comparing both existing and proposed system, the proposed system helps to reduce the deviation of data and helps to early analysis of disaster condition.
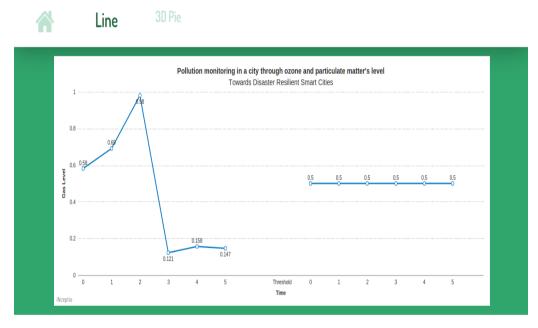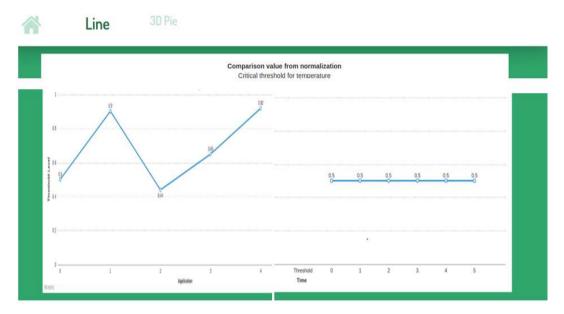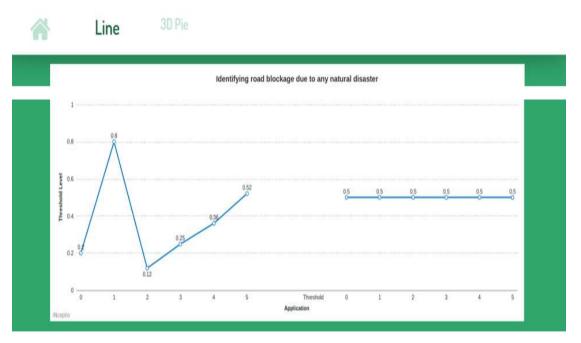

Figure 7. Result of Pollution level monitoring system

Figure 8. Result of fire level monitoring System

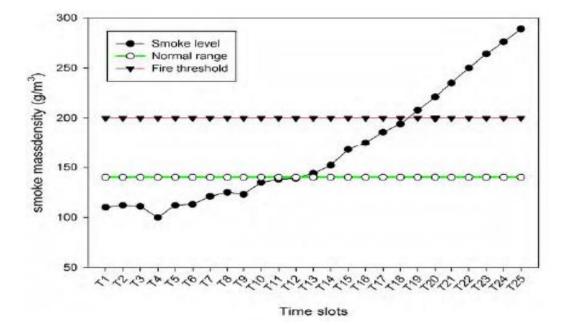**Figure 9.** Result of road blockage monitoring system

Figure 10. Result of pollution level monitoring system of existing system. Courtesy ("Towards Disaster Resilient Smart Cities: Can Internet of Things and Big Data Analytics Be the Game Changers?", IEEE trans. on special section on urban computing and intelligence**.)**

Figure 11. Result of fire level monitoring system of existing system. Courtesy ("Towards Disaster Resilient Smart Cities: Can Internet of Things and Big Data Analytics Be the Game Changers?", IEEE trans. on special section on urban computing and intelligence.)
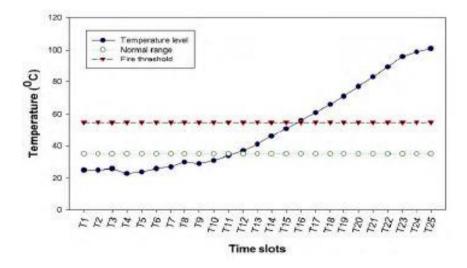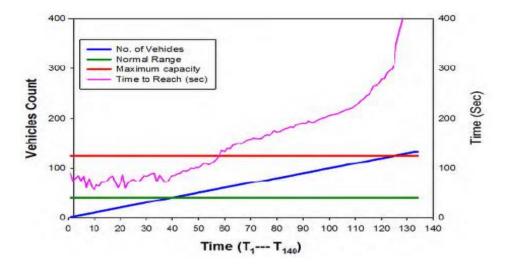
Figure 12. Result of blockage level monitoring system of existing system. Courtesy ("Towards Disaster Resilient Smart Cities: Can Internet of Things and Big Data Analytics Be the Game Changers?", IEEE trans on special section on urban computing and intelligence.)

In normal mean normalization, only the average value of readings will be taken. But, in this reading there can be many unwanted values such as outliers. Therefore, the average value that we take as threshold value will not be the accurate and after data interpretation will also be wrong, which leads to false analysis of disaster condition. This can confuse public services and people with false information. So, with all these information it is evident about the advantages of Z-Score normalization over normal mean normalization for the disaster management system using BDA and IoT.

5. Conclusion
Collaboration of IoT and Big Data Analysis in Disaster management environment is an effective technique which can be practiced for ensuring automated and efficient management with less time. This reference also proves how fog computing is efficient comparing to cloud computing in the case of sudden action like when disaster occur. Also deals with fully automated evacuation technique to be used by considering humanitarian aspects. Now a days every year we are facing different kind of disasters and their after effects. So it is very necessary to implement efficient disaster management technique to reduce the intensity and after effects of it and also for creating an efficient disaster alerting message before it threatens. Also, it shows how efficient clustering system is helping for data analysis. Therefore, based on the inference obtained from the results, it clearly explains how BDA and IoT based disaster management system is necessary to implement in our society mainly in smart cities.

References
[1] Syed Attiquue Shah , Mazhar Rathore , Sadok Ben Yahia and Dirk Draheim 2019, Towards disaster resilient smart cities: can internet of things and big data analytics be the game changers, *Special Section on Urban Computing and Intelligence,* **8,** 91885-91903.
[2] Dursen Safer Seker, 2019, The rising role of big data analytics and IoT in disaster management: recent advances, taxonomy and prospects, *Special Section on Urban Computing and Intelligence*, **7.3**, 54595-54614.

[3] JunboWang and Michael Conrad Meyer, 2019, Maximum data-resolution efficiency for fog-computing supported spatial big data processing in disaster scenarios, *Parallel and Distributed System*, **30.8**,1826-1832 .

[4] Yu wang and Michael Conrad Meyer, 2018, Real-time delay minimization for data processing in wirelessly networked disaster areas, *Special Section on Advanced Big data analysis for Vehicular Social Network,* **7.9**, 2928-2937.

[5] Sandeep K Sooda, Rajinder Sandhua B, Karan Singlaa and Victor Chang, 2018, IoT and Big data based smart flood management framework, *Sustainable Computing Informatics and System,* **20**, 1194-1205.

[6] Xiaolong Xu , Lei Zhang, Stelios Sotiriadis, Eleana Asimakopoulou, Maozhen Li, and Nik Bessis, 2018, CLOTHO: A large-scale Internet of Things-based crowd evacuation planning system for disaster management, *IoT Management,* **5.5**, 3559-3568.

[7] Md Kamruzzaman, Nurul I Sarkar, Jairo Gutierrez and Sayan Kumar Ray, 2017, A study of IoT based post-disaster management, *Information Networking (ICOIN)*, 1221-1226.

[8] Muhammad Arslan, Ana-Maria Roxin,Cristophe Cruz and Dominique Ginnah, 2017, Review of application of big data analysis for disaster management*, 13*[th] *International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*.

[9] Azzedine Boukerche and Rodolpho W L Coutinho, 2018, Smart disaster detection and responses system for smart cities*, Symposium on Computers and Communications*.

[10] Sunil kumar and Maninder Singh, 2019, A novel clustering technique for efficient clustering of big data in hadoop system, *Big Data Analysis and Mining Approach* **2.4** 240-247.