



Wearable IoT sensor based healthcare system for identifying and controlling chikungunya virus



Sandeep K. Sood, Isha Mahajan*

Department of Computer Science and Engineering, GNDU, Regional Campus, Gurdaspur, Punjab, India

ARTICLE INFO

Article history:

Received 2 December 2016

Received in revised form 5 May 2017

Accepted 31 May 2017

Available online xxx

Keywords:

Chikungunya

Fog computing

Cloud computing

Fuzzy-C means

Social network analysis

ABSTRACT

Chikungunya is a vector borne disease that spreads quickly in geographically affected areas. Its outbreak results in acute illness that may lead to chronic phase. Chikungunya virus (CHV) diagnosis solutions are not easily accessible and affordable in developing countries. Also old approaches are very slow in identifying and controlling the spread of CHV outbreak. The sudden development and advancement of wearable internet of things (IoT) sensors, fog computing, mobile technology, cloud computing and better internet coverage have enhanced the quality of remote healthcare services. IoT assisted fog health monitoring system can be used to identify possibly infected users from CHV in an early phase of their illness so that the outbreak of CHV can be controlled. Fog computing provides many benefits such as low latency, minimum response time, high mobility, enhanced service quality, location awareness and notification service itself at the edge of the network. In this paper, IoT and fog based healthcare system is proposed to identify and control the outbreak of CHV. Fuzzy-C means (FCM) is used to diagnose the possibly infected users and immediately generate diagnostic and emergency alerts to users from fog layer. Furthermore on cloud server, social network analysis (SNA) is used to represent the state of CHV outbreak. Outbreak role index is calculated from SNA graph which represents the probability of any user to receive or spread the infection. It also generates warning alerts to government and healthcare agencies to control the outbreak of CHV in risk prone or infected regions. The experimental results highlight the advantages of using both fog computing and cloud computing services together for achieving network bandwidth efficiency, high quality of service and minimum response time in generation of real time notification as compared to a cloud only model.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Chikungunya is the most rapidly spreading infectious virus which is resulting in a new threat to the public health globally. It is transmitted into the human body from two types of infected mosquitoes i.e. *Aedes aegypti* and *Aedes albopictis*. This virus can spread in three ways: (1) Infectious female mosquitoes can transmit CHV to healthy humans. (2) Infectious humans can transmit it to healthy female mosquitoes. (3) Infectious female mosquitoes can lay infectious eggs. It was first identified in 1952 in Tanzania. According to World Health Organization, 6,93,489 suspected cases and 37,480 confirmed cases of chikungunya were reported to the Pan American Health Organization (PAHO) regional office. Out of these, 2,65,000 suspected cases were from Brazil,

1700 cases from Kenya and 19,000 were from Bolivia and Colombia respectively [1]. Henceforth, chikungunya remains a threat for the region. CHV has been identified in 60 countries of Africa, Asia, America and Europe. The symptoms of chikungunya are similar to dengue and zika virus which include most common symptoms like high fever, joint pain, headache, muscle pain, red eyes, nausea and rashes [2]. To differentiate CHV from dengue and zika virus, symptoms based comparison is represented in Table 1. In Table 1 ‘++’ indicates the major symptoms of dengue, zika and chikungunya virus, ‘+’ represents the symptoms which are not so major but they are still significant indicators of these viruses, ‘+’ represents symptoms which are very less frequent and they may or may not appear in person suffering from these viruses and ‘–’ indicates the symptoms which are absent in the person suffering from corresponding disease.

Due to the substantial spread of the CHV, its rapid and early diagnosis is the most significant challenge for government aided health care agencies and developing countries. The most important step to manage CHV is to reduce its rate of infection and

* Corresponding author.

E-mail addresses: san1198@gmail.com (S.K. Sood), isha_3883@yahoo.co.in (I. Mahajan).

migrated onto the cloud platform, which can provide services remotely in a real time [3]. However, there exist many challenges in these systems such as latency issue, location awareness and transmission of big data. As big data is transmitted over a network, the chance of error is high because data transmission latency and packet dropping possibility are proportional to the volume of transmitted data. In case of emergency, single error in data transmission causes inaccurate diagnosis as well as delay in alerting the user which affects the human life. Therefore, there is a need to reduce the data transmission between mobile end users and cloud servers. Healthcare applications also require immediate analysis of data and real time decision without any delay which is not possible in cloud computing. So, a fog layer is included in between the cloud server and mobile end users. First layer consists of IoT based sensors to generate data for further processing on second layer. Second layer is fog layer to pre-process the data which is generated by IoT devices in first layer for real time decisions. Third layer is the cloud layer for storing and processing the data which cannot be processed by fog layer. Use of fog computing in IoT based healthcare system facilitates various benefits such as effective utilization of resources, quality of service assurance, availability of medical data on time and immediate emergency notification [4]. These major benefits encourage the use of fog as a middle layer between IoT sensors and cloud computing layer to identify CHV infected users and also to generate real time notifications to end users in time.

The proposed system addresses the limitations faced in CHV identification and outbreak prevention. Its primary objectives are:

- To design a fog based system for providing remote diagnosis of CHV based on user's health symptoms and surrounding environment conditions.
- To generate immediate diagnostic and emergency alerts to users to enable them to take appropriate action on time.
- To calculate ORI from SNA graph to represent the probability in spreading or receiving the infection.
- To generate warning alerts to government and healthcare agencies to control the outbreak in risk prone and infected regions.
- To protect the confidentiality of sensitive information of users to avoid unauthorized access of data.

To achieve these objectives, fog based health monitoring system is proposed for real time monitoring and analysis of user's health statistics and related events such as health data, location based data, medicinal data, environmental data and meteorological data.

The accumulated data related to different events are integrated into a common format and send to fog layer rather than being sent to the cloud layer for processing in order to obtain fast and high rate service. In fog layer, FCM is used to classify user into possibly infected or uninfected category using health symptoms of the user. Once the user is diagnosed, the fog layer will provide diagnostic and emergency alert messages to user so as to take necessary action on time and provide preventive measures. On cloud layer, ORI is calculated for each user to represent his/her probability in spreading or receiving the infection. Early warning alerts from cloud layer are also generated to uninfected citizens visiting or living in risk prone or infected regions to take immediate precautions to prevent the outbreak of this CHV.

This paper is divided into various sections. Section 2 describes IoT and fog based intelligent system in detail. An experimental case study is performed with the various results in Section 3. Finally, Section 4 concludes the paper.

2. Proposed system

The proposed system shown in Fig. 1 for diagnosing CHV is comprised of three layers, namely, wearable IoT sensor layer, fog layer and cloud layer. Wearable IoT sensor layer collects data in real time from various health sensors, location sensors, drug sensors, environmental sensors and meteorological sensors. The acquired data is transmitted to the fog layer for real time processing and diagnosing possibly infected users from CHV. After diagnosing the CHV, fog layer immediately generates alerts to the user's mobile phones to take precautionary measures on time. At the same time, analysis results and compiled medical information of each user are also stored on cloud layer to calculate ORI of each user to represent his/her probability in spreading and receiving the infection. Cloud layer also generates warning alerts to uninfected citizens visiting or living in risk prone or infected regions to take immediate precautions or necessary actions for user wellness and helps the government agencies to control CHV in such risk prone and infected regions efficiently. The detailed description of each layer of the proposed system is as follows.

2.1. Wearable IoT sensor layer

IoT sensor layer is responsible for collecting data about health related symptoms and various events inside and surrounding environment related to the user. The collected data include health data, environmental data, medicinal data, location based data and meteorological data. Data are collected from the wireless hardware

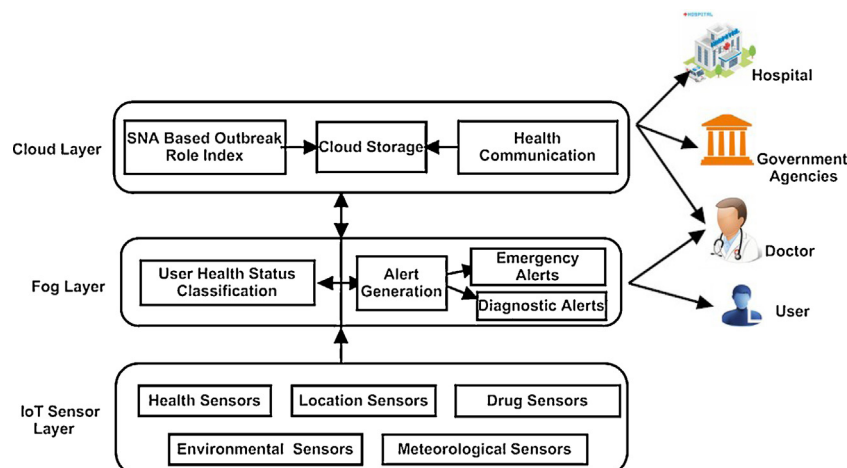


Fig. 1. Fog based proposed system to identify and control chikungunya virus.

devices embedded into the user's body, inside and surrounding places of user. These devices have the capability of sensing and transmitting data in real time. Table 3 provides an overview of various sensors to monitor user's health signs and related events. Table 4 describes the sensitivity, selectivity, reliability and dynamic range of each and every sensor in detecting CHV. Value of all these attributes is transmitted to fog layer to classify user as possibly infected or uninfected. The detailed description of dataset is described below:

- Health data:** Health data set comprises of vital signs of user health. This type of dataset includes high fever, joint pain, headache, body pain, red eyes, rashes on the body, nausea, muscle pain and vomiting. Various health sensors are embedded in the user's body to capture such kind of data.
- Environmental data:** This dataset considers information about users nearby mosquito dense sites and mosquito breeding sites. This information is captured continuously through wireless mosquito sensors which are placed at different locations. Air temperature, humidity and carbon dioxide values around standing water of ponds, lake, dam, river, tube well, refrigerator, cooler, and air conditioner are also measured by sensors to evaluate the environment condition under which the mosquitoes may lay eggs. Sewer quality and water flow rate are also sensed by various sensors embedded at different places of home and surrounding areas.
- Medicinal data:** Medicinal data include medication procedure of the user/patient. It comprises of medicine name, medicine form, quality and ingestion time of the medicine which are included in the medication procedure of the user/patient. This data is acquired through RFID tags. Time is trivially annexed by gateway to this dataset.
- Location data:** It includes location of possibly infected users, location of mosquito dense sites and breeding sites. These locations are sensed by Global Positioning System (GPS) sensors to get travel history of each user during CHV infection.
- Meteorological data:** It includes climate data such as maximum temperature, minimum temperature, amount of rainfall and humidity. These are sensed by various climate sensors.

2.2. Fog computing layer

Fog computing layer acts as a bridge between IoT sensors and cloud computing layer. It is used for real time processing and analysis of accumulated data from IoT based sensors. It sends immediately real time notification or alert to user about current category of user as possibly infected or uninfected. This system is further connected to cloud layer for storing, analyzing results and compiled medical record of each user. Fog computing layer consists of two components such as user health status classification, and alert generation. The detailed description of various components is as follows.

Table 3
Classification of datasets.

S. No.	Data set	Attributes	Wireless sensors
1	Health data	Severe fever, body pain, rashes on body, conjunctivitis, headache, nausea, vomiting, muscle pain	Body sensors
2	Environmental data	Water quality, air temperature, humidity, carbon dioxide, mosquito density	Water quality detector sensor, climate sensor, mosquito sensor
3	Medicinal data	Strength, type, form, proportion	RFID tag
4	Location data	Location of mosquito dense sites, mosquito breeding sites, time	GPS sensor
5	Meteorological data	Maximum temperature, minimum temperature, rainfall, humidity	Climate detector sensor

2.2.1. User health status classification

Classification is an important tool for taking decision in various medical diagnoses. This component provides an initial diagnosis to users by classifying them depending upon their CHV attribute values using FCM classifier as possibly infected and uninfected. FCM [23] is the technique of clustering of data in which a dataset is organized into n clusters such that the each data point in the dataset belongs to each cluster to a definite amount of degree. The degree of the belonging or membership of certain data point in the cluster is high when it lies close to the center of that cluster and the degree of the membership is low when it lies far away from the center of a cluster. FCM repeatedly computes the centroids of each cluster and the fuzzy pseudo-partition until the partition does not change with time. A set of ten data points are denoted as $Y = \{Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}\}$ where these $Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7, Y_8, Y_9, Y_{10}$ refer to the attributes i.e. fever, joint pain, headache, muscle pain, red eyes, skin rash, nausea, sore throat, itching and fatigue. Two fuzzy clusters, S_1 , and S_2 are subset of all possible fuzzy subsets of Y which are used to classify users by grouping their dataset into two clusters in which one cluster contains uninfected users and second cluster contains possibly infected users. The main objective of FCM is to minimize errors in the objective function. Mathematical form of objective function is shown in Eq. (1).

$$J(U, A, Y) = \sum_{i=1}^2 \sum_{j=1}^{10} u_{ij}^m d_{ij}^2(y_j, a_i) \quad (1)$$

$$J(U, A, Y) = \sum_{i=1}^2 \sum_{j=1}^{10} u_{ij}^m \|y_j - a_i\|^2$$

where U is the partition matrix such that $U = u_{ij}$, $i = 1, 2$ and $j = 1, 2, \dots, 10$. u_{ij} is the membership value of the j th data to the i th cluster. Y is the set of data points. A is the set of cluster centers $A = a_i$. m is the fuzzy strength parameter and is generally chosen as 2 for optimal statistical results. It controls the fuzziness of the system. $\|y_j - a_i\|^2$ is the Euclidean distance between data points y_j and cluster centers a_i of the i th cluster. Following three conditions of membership value must satisfy:

- Range of membership values between 0 and 1 is given in Eq. (2)

$$u_{ij} \in [0, 1], \quad \forall i, j \quad (2)$$

- The summation of the membership values of each data point must be 1 as given in Eq. (3)

$$\sum_{i=1}^2 u_{ij} = 1, \quad \forall j \quad (3)$$

Table 4

Description of sensors.

Parameters	Name of sensor	Selectivity	Sensitivity	Dynamic range	Reliability
Environmental attributes	Climate Meter PCE-FWS 20	(a) Measures temperature (b) Relative humidity, (c) Precipitation, atmospheric pressure (d) Wind speed and wind direction (e) Storable measurement values (with date and time) (f) Alarm function	2.6–4.1 °C	(a) Measurement range: 0 °C+...+60/+32 °F +...+140 °F (b) Humidity: 1...99% (c) Atmospheric pressure: 919...1080 hPa	±3% F.S. (<20 m/s)/±4% F.S. (>20 m/s)
Severe fever	Caregiver Touch Free Thermometer	Measure forehead temperature in adults, infants and children without contact which reduces the risk of cross contamination	Highly accurate (no patient disturbance)	Measure in 1–2 s	±0.5 °F
Body pain, abdominal pain and muscle pain	Neat-O	Easy to use by put the device on their skin near their pain	0.2–1.2 m	1–10 (pain score)	140 DB
GPS	LK209B Magnet Vehicle GPS	(a) Real time tracking (b) Auto tracking (c) Low battery alert (d) Blind area tracking (e) Over speed alert (f) History trace checking	–159 dBm	(a) Storage temperature: –40 °C to +85 °C (b) Operation temp.: –20 °C to +55 °C %–95% non-condensing	5 m
Mosquito sensor	BG Counter	(a) Automatic start-up (b) Fan and CO ₂ flow check (c) Cellular connection check (d) Counter check	–	From anywhere	90%

- The summation of all membership values in each cluster must be lesser than number of data points (N) as given in Eq. (4).

$$0 < \sum_{i=1}^8 u_{ij} < N \quad (4)$$

The details of the FCM algorithm are presented in Algorithm 1. An Algorithm 2 is designed to evaluate the category of the user using FCM classification algorithm.

Algorithm 1. Fuzzy C-means algorithm.

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|y_j - a_i\|}{\|y_j - a_k\|} \right)^{\frac{2}{m-1}}}$$

- 1: Determine the number of clusters and also set the value for fuzzifier constant. Threshold value ($\varepsilon > 0$) of the termination condition is also to be assigned which is to be used for process termination.
- 2: Initialize the membership matrix
 $U = \{u_{ij}\} c \times n$ where c is number of clusters and n is the number of data points.
- 3: Calculate fuzzy cluster centers a_i using following equation and k is the iteration number
 $a_i^k = \frac{\sum_{j=1}^n (u_{ij}^{(k-1)})^m y_j}{\sum_{j=1}^n (u_{ij}^{(k-1)})^m}$ where $i = 1, 2, \dots, c$
- 4: Update membership value u_{ij} with a_i^k using following equation
 $u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|y_j - a_i\|}{\|y_j - a_k\|} \right)^{\frac{2}{m-1}}}$
- 5: If $\|U^{k+1} - U^k\| < \varepsilon$, stop otherwise increment k to follow step 3. This process for solving membership degree and cluster centers continuous until termination condition is satisfied.

Algorithm 2. To evaluate the health symptoms of user to classify category.

Input: CHV health attributes and identification number of user
Output: Classified category of a user based on health attributes

Step 1. Get CHV health attributes and identification number of user.

Step 2. If identification number is already present in database

Step 2.1 update the database with newly entered data.

Step 3. Else

Step 3.1 Create a new record with identification number of the user and store health attributes.

Step 4. Execute FCM to predict the category of the user.

Step 5. Store classified category of user in the database with corresponding identification number.

Step 6. Exit

Algorithm 3. To generate emergency alerts to users, doctors and healthcare professionals.

Input: Current classified category of user, probability of various events and predefined threshold value

Step 1. Get classified category of user, health attributes and events of user of current time stamp.

Step 2. If classified category = possibly infected category

Step 2.1 Calculate Sensitivity factor of possibly infected user and probability of various events of current time stamp.

Step 3. If (Sensitivity Factor > Predefined Threshold)

Step 3.1. User is in unsafe state and immediate emergency alert is generated on user's mobile phone.

Step 4. Else

Step 4.1. user state is safe and no alert is generated to user.

Step 5. Exit

2.2.2. Alert generation component

Alert generation component is responsible for immediately sending alert messages on user's mobile. It comprises two alerts namely, diagnostic and emergency alerts. Diagnostic alerts are generated for all users either possibly infected or uninfected. Whereas, emergency alerts are generated only for possibly infected

users in an emergency condition. Detailed working of diagnostic and emergency alerts is as follows:

2.2.2.1. Diagnostic and emergency alerts. Once users are diagnosed, this component will send immediate diagnostic alert message on user's mobile. Diagnostic alert message contains diagnosed category of user as possibly infected or uninfected of CHV. If user is classified as possibly infected from CHV, infected user is continuously monitored with various events such as medicinal, environmental, climate and location. In this case, sensitivity factor with respect to time of various events is calculated as:

$$\text{Sensitivity Factor} = P\left(\frac{S}{(E_1 \cup E_2 \dots \cup E_n)}\right)$$

where S denotes the user's health category infected as CHV. $\{E_1, E_2, \dots, E_n\}$ represents the probability of occurrence of n events in current time. If the sensitivity factor of infected user is greater than predefined threshold then automatically alert is generated to concerned user for taking necessary actions in the direction of wellness. It also provides information about nearby hospitals to the users depending upon the location of user and shares his/her status of infection with nearby doctors to immediately perform test for CHV. Threshold value of each patient is predefined based on the health condition and health history. Otherwise, no alert is generated on user's mobile if sensitivity factor is less than threshold. These alerts are also stored on cloud storage to share with doctors and healthcare professionals during emergency state of the patient. Therefore, sensitivity factor based alert generation enhances the decision making capabilities of doctor in a time sensitive manner and reduces the number of false alert generation. Overall working of emergency alert generation is depicted in Algorithm 3.

2.3. Cloud layer

Cloud layer is responsible for storing and processing the data which cannot be processed by fog layer. It consists of four components namely; cloud storage, SNA based outbreak role index, information protection and health communication. The following subsections will discuss in details about these components.

2.3.1. Cloud storage

The objective of cloud storage is to store information related to possibly infected users or uninfected users, location of mosquito dense and breeding sites. It consists of huge amounts of storage to store analysis results and compiled medical information of each user and securely share among authorized medical staff, users, pharmacies, hospitals and healthcare professionals. Diagnostic and emergency alert messages generated from fog layer are also stored on cloud storage for further analysis by experts to take immediate action and provide precautions in case of emergency. Travelling information of infected and uninfected users is stored on cloud storage to create or update global SNA graph with time and

generate warning alerts accordingly. Government aided healthcare centers can also upload data as well as any information regarding first aid, free camps, etc. to control the outbreak of CHV.

2.3.2. SNA based outbreak role index

CHV spread over large distance or from one region to another region by humans because flight range of mosquitoes is very limited. Both humans and mosquitoes are classified as healthy and infectious. This virus can be spread in three ways: (1) Infectious female mosquitoes can transmit CHV to healthy human. (2) Infectious human can transmit it to healthy female mosquitoes. (3) Infectious female mosquitoes can lay infectious eggs. There are various cases to spread virus from one region to another region or in the same region:

Case 1. Human infection in region 'i' take place when infected mosquitoes living in region 'i' bites a healthy person who is:

1. Resident in region 'i' and present in region 'i'.
2. Resident in region 'j' and present in region 'i'.

Case 2. Mosquito infection in region 'i' take place when a susceptible mosquito bites an infected person who is:

1. Resident in region 'i' and present in region 'i'.
2. Resident in region 'j' and present in region 'i'.

Algorithm 4. To create global SNA graph.

Input: Infected or Uninfected user, their resident and travelling history
Output: Newly or updated global SNA graph
Step 1. Get classified category of user, resident and travelling locations.
Step 2. If user classified category = Possibly infected Then
 Step 2.1 Create two nodes one for user and second for his/her residence with Red color.
 Step 2.2 Get travelling locations of user.
Step 3. Else
 Step 3.1 Create a new node of user's residence with Green color.
 Step 3.2 Get travelling locations of user
Step 4. For $i = 1$ to n // n is the number of regions in SNA graph
 Step 4.1 if travelling location $[i]$ is already in the graph.
 Step 4.1.1 create a new edge between travelling location $[i]$ and user.
 Update color of regions according to the category of user
 Step 4.2 Else
 Step 4.2.1 create a new node with travelling location $[i]$
 Step 4.2.2 create a edge between travelling location $[i]$ and user. Update color of region according to the category of user.
Step 5. Exit

The objective of this component is to identify infected and risk prone regions so that the spread of CHV infection can be controlled by sending alert messages and infection control suggestions to the government agencies and to the people residing or visiting in these regions. To identify the infected and risk prone regions, SNA is used [24]. It provides various tools and techniques to analyze and extract results from complex graphs. Tools and techniques of SNA are useful if a social network graph has been created efficiently. So,

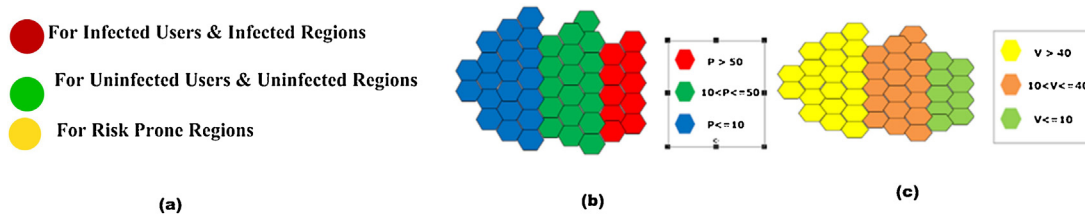


Fig. 2. (a) Coloring scheme for users and regions in SNA graph. (b) Hexagonal representation of CHV infected regions. (c) Hexagonal representation for CHV risk prone regions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Algorithm 4 provides step by step computation for creating global SNA graph. In this graph, once user is screened as possibly infected by classification component, his/her travelling locations are traced through GPS enabled mobile phones. These locations are added in the global SNA graph and make edges between these locations and user. It is created and updated on time. Different color schemes are also used for different regions as well as different users while creating SNA graph as shown in Fig. 2(a). In SNA graph, red color for infected regions and possibly infected users, green color for uninfected regions and uninfected users, and yellow color for risk prone regions. Infected regions are those regions where one or more individuals are infected. Uninfected regions are those where no one is infected from CHV. Risk prone regions are those where one or more infected individuals are visited and at high risk to spread the CHV infection.

Many important outputs can be drawn from global SNA graph which will help government health agencies to control the CHV outbreak in infected regions. There are several outbreak metrics which can be computed from SNA graph. These metrics are explained below:

Metric 1: Outbreak role index (ORI).

ORI calculates the probability of any user to receive or spread the infection. It is calculated from SNA graph.

Definition 1. ORI for infected user and infected regions states the ratio of number of infected regions visited by infected user to the total number of regions visited by infected user.

$$\text{ORI (Infected User and Infected Regions)} = \frac{\text{of infected regions visited by infected user}}{\text{Total number of regions visited by infected user}}$$

Definition 2. ORI for infected user and uninfected regions states the ratio of number of uninfected regions visited by infected user to the total number of regions visited by infected user.

$$\text{ORI (Infected User and Uninfected Regions)} = \frac{\text{of uninfected regions visited by infected user}}{\text{Total number of regions visited by infected user}}$$

Definition 3. ORI for uninfected user and infected regions states the ratio of number of infected regions visited by uninfected user to the total number of regions visited by uninfected user.

$$\text{ORI (Uninfected User and Infected Regions)} = \frac{\text{of infected regions visited by uninfected user}}{\text{Total number of regions visited by uninfected user}}$$

There are some cases which will require special attention of healthcare agencies.

1. If ORI is high in case of infected user and uninfected regions, it means infected user visited many uninfected regions. So this user poses high probability of spreading the CHV infection. So these users should be quarantined as early as possible. Warning

alerts should be sent to government and nearby hospitals of risk prone regions to take immediate action to control the outbreak of CHV.

2. If ORI is high in case of uninfected user and infected regions, it means uninfected user visited many infected regions and at very high risk to get infection. So immediately alert should be sent to these types of user.
3. If ORI is high in case of infected user and infected regions, it means infected users and infected regions should be examined as early as possible. Warning alerts should be sent on user's mobiles and to government agencies to take action to control the outbreak.

Metric 2: Relative score of each region.

Relative score of each region will help to identify critical regions from the global SNA graph. Appropriate alerts can be generated based on the score values. It will help the government agencies to identify critical regions and eliminate all kind of travel to and from those regions. Government agencies can take preventive measures to control the CHV outbreak of that critical region. Relative score S_i of any region i can be computed as

$$S_i = 1/\lambda \sum_{j \in U(i)} s_j = \frac{1}{\lambda} \sum_{i=1}^n A_{ij} * S_j$$

where $U(i)$ is the set of infected users nodes which are connected to the 'ith' region. ' λ ' is the Eigen constant and A_{ij} is the adjacency matrix. A region with high score should be quarantined as early as possible.

From these metrics, risk level of CHV in possibly infected and risk prone regions are identified. According to the risk level of CHV infection, these regions are mapped on Google map web service so that people living or visiting in these regions can be alerted on time and suggestions to control infection can be issued to them. Hexagonal based mapping is used to represent infected and risk prone regions as shown in Fig. 2(b) and (c). A distinct color scheme is used to efficaciously represent these regions based on the density of possibly infected users as well as density of infected visitors visiting in any particular region as given in Table 5. The regions of new infected users and their travelling history are automatically detected by the system and Google map is updated accordingly. Algorithm 5 is designed to make hexagonal structure dynamic and plot on Google maps. The representation of infected regions and risk prone regions on Google map help the government healthcare agencies and uninfected users to control the epidemic of CHV infection.

Algorithm 5. To map infected and risk prone regions on Google map.

Input: SNA graph

Output: Create or update Google map

Step 1. Identify possibly infected and risk prone regions from SNA graph.

Step 2. For every possibly infected region.

Step 2.1. Calculate whole population and infected users in hexagonal structure of that region.

Step 2.2. Increment the density of hexagonal structure.

Table 5 should appear before Algorithm 5..because we have mentioned Table 5 before algorithm 5 in written text. But in proof pdf Algorithm 5 is appearing before Table 5 please correct it if possibleTable 5

Different risk level of CHV for infected and risk prone regions.

Levels of CHV infection for infected regions		Levels of CHV for risk prone regions	
Infected population (P) in %	Risk level (color)	No. of infected visitors (V) in %	Risk level
$P > 50$	High (red)	$V > 40$	High (dark yellow)
$10 < P \leq 50$	Medium (green)	$10 < V \leq 40$	Medium (orange)
$P \leq 10$	Low (blue)	$V \leq 10$	Low (light green)

- Step 2.3. Update hexagonal structure's color based on computed hexagonal's density, represented in Table 5.
- Step 2.4. Plot hexagonal on Google map.
- Step 3. For every risk prone region.
- Step 3.1. Calculate total number of infected visitors visited in risk prone region.
- Step 3.2. Increment the density of hexagonal.
- Step 3.3. Update hexagonal structure's color based on computed hexagonal's density, represented in Table 5.
- Step 3.4. Plot hexagonal on Google map.
- Step 4. Exit

2.3.3. Information protection

Cloud storage component stores personal and health details of the user. It contains some highly sensitive attributes which cannot be shared with all. Even an unintentional disclosure of information of such attributes to an unauthorized person may lead to a panic situation among individuals. The proposed system uses two stage information protection mechanism i.e. information fragmentation and key sharing scheme to prevent the unauthorized access of data. Both stages are explained below:

2.3.3.1. Information fragmentation. In this stage, data collected from user is fragmented into three fragments of varying levels of security i.e. level 1, level 2 and level 3. Level 1 is extremely confidential that contains personal data such as age, name, gender, mobile number and residential address. Level 2 consists of medium level of information of environmental attributes and meteorological attributes. Level 3 is least sensitive level of information that contains CHV symptoms related attributes. Even if someone is able to get the information of level 3, the exact identity of the user can not be ascertained. A knowledge of all three levels of fragments is required for retrieving exact identity of user. In an order to resolve the problem of protecting the user's highly sensitive attributes, collected information from each individual should be converted from data table to appropriate data fragments. A data table 'D' having a number of attributes ($a_1, a_2, a_3 \dots a_n$) of user ' U_i ' and divided into three fragments of different levels of security. To correctly fragment the data table, the following constraints should be met:

$$F_1 \cup F_2 \cup F_3 = D$$

$$F_i \cap F_j = \emptyset, i \neq j \text{ and } F_i, F_j \in F$$

2.3.3.2. Secret sharing mechanism. A data table 'D' contains sensitive attributes i.e. a_s should be protected from unauthorized users. So, key sharing mechanism is used to distribute the values of highly sensitive attribute into 'n' parts i.e. $v_1, v_2, v_3, \dots, v_n$. These 'n' pieces are stored on different secure cloud servers. Choose $k-1$ coefficients a_1, a_2, \dots, a_{k-1} and assign key value a_s to the coefficient a_0 . Polynomial of degree $(k-1)$ is

represented in Eq. (5)

$$F(x) = a_0 + a_1x + a_2x^2 + \dots + a_{k-1}x_{k-1} \quad (5)$$

Choose secret keys $X = x_1, x_2, \dots, x_n$ which are randomly chosen values corresponding to each cloud server. System computes each cloud server share by substitute the value of $x_i \in X$ and values of the coefficients a_1, a_2, \dots, a_{k-1} in Eq. (5) as $F(x_i)$ and store it to the corresponding cloud server. To recreate the original value of highly sensitive attribute from key value of shares, knowledge of any 'k' pieces where $k \leq n$ along with secret information i.e. x_1, x_2, \dots, x_n is required. The secret information is only stored on the trusted server, thus only trusted server can recreate the original value of highly sensitive attribute after retrieving at least 'k' shares from any 'k' cloud servers. Knowledge of any 'k-1' or less pieces is not enough to recreate the original value of attribute even if secret keys X is known to cloud servers.

2.3.4. Health communication component

This component is used to control the spread of CHV. System will generate alert messages and health education to user via instant message or e-mail to prevent the growth of CHV outbreak. Alert messages and suggestions regarding health information can be sent repetitively to user's mobile phone to improve knowledge of preventive measures which may lead to reduction in CHV risk. Alert messages are also sent to nearby hospitals as well as to government healthcare agencies depending upon the GPS location of the patient's mobile phone to provide first aid, organize free camps for general public. Tables 6 and 7 shows the suggestions and preventive measures for uninfected and infected users.

3. Experimental setup and performance evaluation

This section represents the experimental setup and discusses the performance evaluation of the proposed system. Due to non-availability of symptoms based data set of CHV, data are systematically generated by consulting specialists. The detail of data generation is given in Section 3.1. Experimental section is divided into five sections i.e. Generation of symptoms based datasets, data transmission to fog layer, efficiency of alert generation, evaluation of SNA graph based risk assessment and comparative analysis of power consumption rate.

3.1. Generation of symptoms based datasets

A thorough search on internet, we were not able to find any CHV database based on the symptoms of patients to test the proposed system. So symptoms based datasets to evaluate proposed system is systematically generated so that no possible case has been left out. Table 8 shows the probabilities of each CHV symptoms to be present in any new generated case while creating dataset for CHV. A real dataset of 5000 cases containing environmental attributes are obtained from [25,26] and monthly climate attributes, namely

Table 6
Suggestions and preventive measures for uninfected users.

S. No.	Suggestions
1	Use of liquids, mats, mosquito repellent creams, and coils
2	Use full sleeve shirts and full pants with socks during the transmission season
3	Use of bed nets while sleeping to prevent mosquito bite
4	Always cover trash or dustbin
5	Clean the water tray of refrigerator and air conditioner regularly
6	All the containers are covered that hold water to prevent mosquito to access the water
7	Avoid to visit near shrubby areas where adult mosquitoes usually rest during the day
8	Remove water from coolers and other small containers at least once in a week
9	Do not play in shorts and half sleeved clothes
10	Door screens and windows do not have any holes. If so, cover those areas to prevent mosquitoes

Table 7

Suggestions and home remedies for infected users.

S. No.	Suggestions
1	Immediate consult with doctor for treatment
2	Use turmeric, chili pepper and garlic in regular diet
3	Eat grapes with cow milk to get quick relief from chikungunya
4	Eat citrus fruits with regular diet
5	Eat raw carrots or drink fresh carrot juice
6	Eat green leafy vegetables
7	Drink plenty of coconut water to have relief from CHV
8	Eat apples or drink fresh apple juice to combat the symptoms of chikungunya
9	Add epsom salt in luke warm bath water having neem leaves
10	Mix castor oil with cinnamon powder and massage the affected joints for a few minutes

Table 8

Probabilities for CHV symptoms.

Primary symptoms	Probabilities	Secondary symptoms	Probabilities
High fever	0.18	Red eyes	0.11
Joint pain	0.15	Nausea	0.09
Skin rash	0.11	Itching	0.08
Headache	0.09	Sore throat	0.07
Muscle pain	0.08	Fatigue	0.05
No symptoms	0.40	No symptoms	0.60

temperature, humidity and rainfall are obtained from [27–29] which are integrated systematically with the generated all possible cases of CHV infection related attributes. All possible cases of CHV infection related attributes are mapped randomly with the dataset of environmental and climate attributes as shown in Algorithm 6.

Algorithm 6. To create synthetic data sets for CHV.

```

Input: Data containing CHV symptoms, datasets of environmental, climate
      attributes and number of distinct cases required.
Output: Generate CHV datasets
Step 1. 'n' be the number of CHV cases initialized with 1.
Step 2. For  $n \leq$  required number of cases
Step 3. Assign values to primary symptoms based on probabilities in Table 8.
Step 4. Assign values to secondary symptoms based on probabilities in Table 8.
Step 5. Create a new case by combining all CHV symptoms values with
      environmental and climate attributes.
Step 6. If new case is already present in database then
      Step 6.1. Discard the new case
Step 7. Else
      Step 7.1. Add the new case and value of generated cases is increased by 1
Step 8. Exit

```

3.2. Data transmission to fog layer

Generated datasets are used to train and test the FCM clustering to categorize users into possibly infected or uninfected. FCM

classifier was implemented on an Intel i7 CPU at 2.50 GHz with 4 Gbytes memory using Matlab running on windows 7 using the dataset of 5000 cases. The dataset is first normalized to transform symptoms values accepted in 'yes' or 'no' into the interval [0,1] according to the strength of each symptom to avoid numerical difficulties during calculation. The classification accuracy of FCM algorithm is tested with fuzzy strength parameter ' m ' which varies in the range [1,2]. For each choice of ' m ', we test the various statistical parameters namely, sensitivity, specificity, accuracy, precision, recall and error rate. Finally, the one with lowest error rates and the highest accuracy are selected as the optimal fuzzy strength parameter. Table 9 summarized the results of classification performance of FCM in terms of accuracy, sensitivity, specificity, precision, recall and error rate. From Table 9, we can see that the results of the classification performance of FCM classifier with different value of ' m ' parameter are very close. But highest accuracy is achieved with ' m ' = 2. The results show that the system is able to efficiently classify the users into possibly infected or uninfected with high value classification accuracy, sensitivity, specificity, precision & recall and low values for error. Fig. 3 represents the comparison of different classification algorithm namely, NN (Neural Network), FKNN (Fuzzy k-Nearest Neighbors) and NB (Naive Bayes) using different number of cases. The accuracy of classification algorithms is shown in Fig. 3(a) which shows that the FCM based classification algorithm performs better than all other algorithms. Fig. 3(b) shows the classification time of different algorithms using different datasets. Fig. 3(c) represents the classification execution time of FCM classifier with different datasets which are transmitted to health fog rather than being sent to the cloud for processing and analysis. The graphical trend shows that proposed system through health fog computing layer identify category of user as possibly infected or uninfected with less time than directly through cloud computing. It is because fog computing provides various resources such as compute, storage and communication in the close proximity of mobile users as compared to cloud computing. It also avoids unnecessary flow of raw information from mobile end user to cloud server while processing and sharing the information.

Table 9

Results of classification performance of FCM using different parameters.

Statistical Parameters	FCM $m = 1.4$ (in %)	FCM $m = 1.6$ (in %)	FCM $m = 1.8$ (in %)	FCM $m = 2.0$ (in %)
Classification accuracy	89.5	90.65	92.98	93.40
Sensitivity	86.7	87.3	88.4	90.45
Specificity	85.2	88.8	90.3	91.23
Precision	85.8	90.0	91.8	91.89
Recall	86.0	89.7	90.4	91.0
Mean absolute error	2.87	1.37	0.34	0.23
Root mean square error	1.30	0.78	0.12	0.10
Relative absolute error	7.89	6.789	4.78	3.78

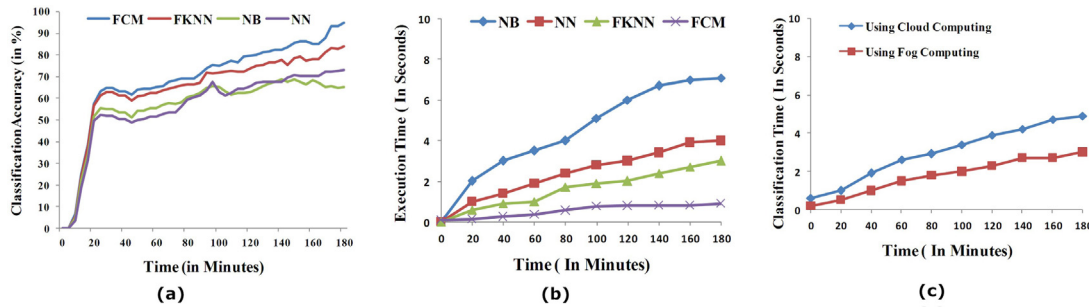


Fig. 3. (a and b) Experimental results: (a) Performance of classification accuracy of different algorithms. (b) Execution time of different algorithms. (c) Total execution time of the FCM classifier using fog computing as compared to cloud computing.

3.3. Efficiency of alert generation from fog layer

Efficiency of alert generation mainly relates to the effectiveness of CHV diagnosing in the generation of timely and true alerts through fog computing. The proposed system is examined for determining the validity of alert generated on user's health condition. The main aim for statistically analyzing the efficiency of alert generation component is to verify the false positive alerts. These alerts are based on total number of alerts and evaluate alert generation in terms of delay to deliver information to the user and doctor. Delay time is calculated between time to deliver alert about the occurrence of event to user and time at which event occurred. As a comparative model, cloud based alert generation to doctors as well as users are compared with fog based alert generation in terms of response time. Fig. 4 represents the results which are depicting the estimated delay time for both cloud and fog based computing. The results show that real time notification from fog system is far efficient as compared to the notification from cloud with low value of delay during the occurrence of any abnormal or emergency situation. Various other statistical parameters are also considered for this purpose, such as sensitivity, specificity, precision, coverage, mean absolute error, root average square error, root relative square error and relative absolute error. All these results are shown in Table 10. Statistical parameters indicate that proposed system is highly accurate and effective as the low value of false positive alerts. Moreover, less error rates also enhance the utility of alert generation method.

3.4. Evaluation of SNA graph based risk assessment

NetLogo [30] is used for creating SNA graph. To evaluate SNA graph based risk assessment of infected and risk prone regions, data related to infected users and their travelling locations during CHV infection are generated over the Pathankot city in India. Two files in the.csv format containing the details of 5000 users and their travelling locations are mapped in the particular block of Pathankot city in the form of hexagonal structure as shown in Fig. 5(a) and (b). Fig. 5(b) shows the routing of a user from location

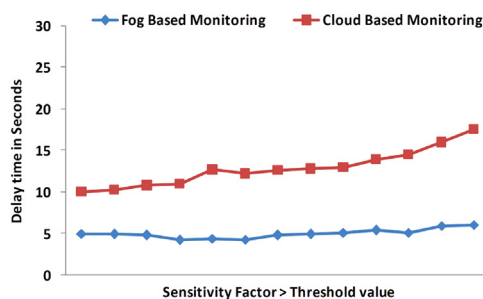


Fig. 4. Efficiency of delay time.

Table 10

Statistical results of alert generation.

Statistical parameters	Values (in %)
Sensitivity	88.4
Specificity	94.5
Precision	91.4
Coverage	96.5
Mean absolute error	2.98
Root average square error	2.50
Root relative square error	34.4
Relative absolute error	7.68
False positive alerts	3.12

A to location B without implementing any routing algorithm. Here, user passes through the infected as well as risk prone regions which increases the risk of catching CHV infection. Alternatively, the user has diverted to the safer path represented by using appropriate routing algorithm as shown in Fig. 5(c). The red line in Fig. 5(b) shows the normal route and purple line in Fig. 5(c) shows the proposed or safer route for the user by Google services.

3.5. Comparative analysis of power consumption rate

Fig. 6 depicts power consumption rate of the proposed system with varying number of datasets. These datasets are transmitted to both proposed system (health fog system and cloud computing system) and other system using only cloud computing technique for comparing the power consumption rate. The comparison results are shown in Fig. 6 which represents the comparison of power consumption rate of both fog computing and cloud computing with respect to different datasets. The trend of graph shows that power consumption is more in case of cloud computing as compared to fog computing. It is because the fog computing provides computation and communication services in the close proximity of end users. It reduces the volume of data transmitted over network which leads to less traffic resulting in low latency rate and ultimately less power consumption. But in case of cloud computing, large amount of data is transmitted to cloud server for processing which leads to higher traffic congestion over network. This has an adverse effect on response time and power consumption rate.

4. Conclusion

CHV is mosquito borne virus that is spread especially by *A. aegypti* and *A. albopictis* mosquito and causes human morbidity and mortality. Due to the advancement in information and communication technology with IoT, fog computing and cloud computing, it is now possible to enhance the level of information acquisition, storage, efficiently diagnose, and provide a global

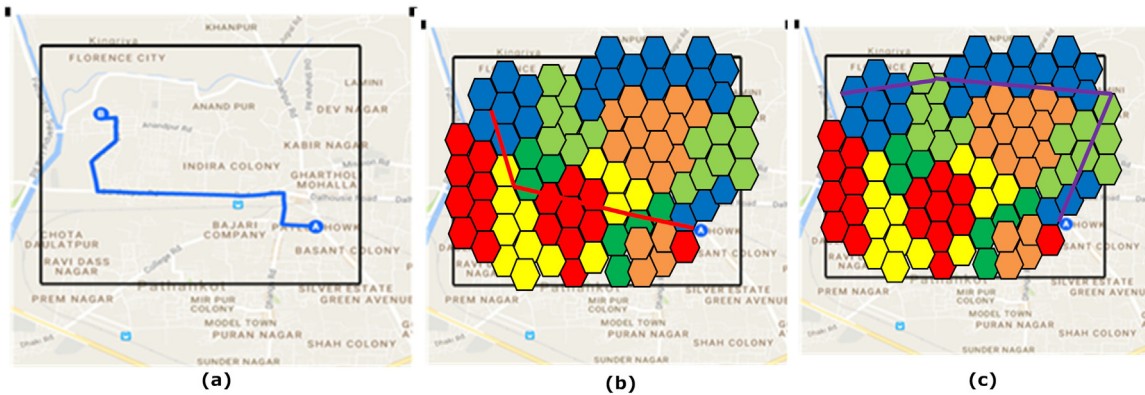


Fig. 5. (a–c) Experimental results: (a) GPS based rerouting of user from location A to location B. (b) Default routing of user from location A to location B. (c) Safe route of the user based on infected and risk prone regions.

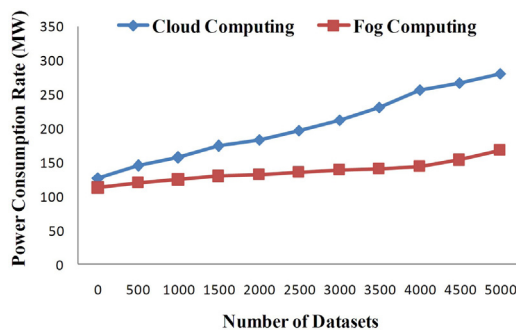


Fig. 6. Comparative analysis of power consumption rate between fog computing and cloud computing.

information access to users, doctors and healthcare professionals. In this paper, fog based intelligent system is proposed to accurately diagnose user as possibly infected or uninfected from CHV by using FCM clustering and immediately send diagnostic alerts to user's mobile from fog system. The main feature of this paper is the use of sensitivity factor with respect to time of various events for determining the health severity. On the basis of health severity, emergency alerts are generated for delivering event information to user's mobile on time through fog network. Furthermore, SNA is used to represent the state of outbreak. ORI is calculated from SNA graph which represents the probability of any user to receive or spread the infection. It also generates warning alerts in due time to uninfected users those are visiting or living in risk prone or infected regions. It will help uninfected residents to take immediate precautions to prevent the outbreak of these viruses and government healthcare agencies to control the problem effectively. The experimental results of proposed system are compared with cloud computing technology. The results reveal that proposed framework achieves bandwidth efficiency, minimum execution time, and minimum delay in generation of real time notifications.

References

- [1] World Health Organization, Chikungunya, (2017) Available from: <http://www.who.int/mediacentre/factsheets/fs327/en/> (accessed 16.04.17).
- [2] Centers for Disease Control and Prevention, Chikungunya Virus, (2017) Available from: <https://www.cdc.gov/chikungunya/symptoms/> (accessed 16.04.17).
- [3] C. He, X. Fan, Y. Li, Toward ubiquitous healthcare services with a novel efficient cloud platform, *IEEE Trans. Biomed. Eng.* 60 (1) (2012) 230–234.
- [4] M. Ahmad, M.B. Amin, S. Hussain, B.H. Kang, T. Cheong, S. Lee, Health fog: a novel framework for health and wellness applications, *J. Supercomput.* (2016), doi:<https://doi.org/10.1007/s11227-016-1634-x>.
- [5] R. Hassan, M.M. Rahman, M. Moniruzzaman, A. Rahman, S. Barua, R. Biswas, P. Biswas, S.G.M. Mowla, M.J. Chowdhury, Chikungunya – an emerging infection in Bangladesh: a case series, *J. Med. Case Rep.* 67 (8) (2014) 1–3.
- [6] F. Gobbi, D. Buonfrate, A. Angheben, M. Degani, Z. Bisoffi, Emergence and surveillance of chikungunya, *Viral Trop. Med.* 2 (1) (2015) 4–12.
- [7] S.C. Weaver, N.L. Forrester, Chikungunya: evolutionary history and recent epidemic spread, *Antivir. Res.* 120 (3) (2015) 32–39.
- [8] X. Liu, P. Stechlinski, Application of control strategies to a seasonal model of chikungunya disease, *Appl. Math. Model.* 39 (12) (2015) 3194–3220.
- [9] E.P. Calvo, F.S. Quete, S. Duran, I. Sandoval, J.E. Castellanos, Easy and inexpensive molecular detection of dengue, chikungunya and zika viruses in febrile patients, *Acta Tropica* 163 (4) (2016) 27–32.
- [10] S.B. Murugan, R. Sathishkumar, Chikungunya infection: a potential re-emerging global threat, *Asian Pac. J. Trop. Med.* 9 (10) (2016) 933–937.
- [11] S.L.B. Silva, S.S.C. Hernandez, E.M. Palacios, J.A.P. Molina, Clinical and differential diagnosis: dengue, chikungunya and zika, *Hosp. Gen.* (2016), doi:<https://doi.org/10.1016/j.hgmx.2016.09.011>.
- [12] K. Pabbaraju, S. Wong, K. Gill, K. Fonseca, G.A. Tipples, R. Tellier, Simultaneous detection of zika, chikungunya and dengue by a multiplex real time RT-PCR assay, *J. Clin. Virol.* 83 (2016) 66–71.
- [13] G. Yang, L. Xie, M. Mantysalo, X. Zhou, Z. Pang, L. Da Xu, S. Kao-Walter, Q. Chen, L.R. Zheng, A health-IoT platform based on the integration of intelligent packaging, unobtrusive bio-sensor, and intelligent medicine box, *IEEE Trans. Ind. Inf.* 10 (4) (2014) 2180–2191.
- [14] B. Xu, L. Da Xu, H. Cai, C. Xie, J. Hu, F. Bu, Ubiquitous data accessing method in IoT-based information system for emergency medical services, *IEEE Trans. Ind. Inf.* 10 (2) (2014) 1578–1586.
- [15] R. Sandhu, S.K. Sood, G. Kaur, An intelligent system for predicting and preventing MERS-Cov infection outbreak, *J. Supercomput.* 72 (8) (2015) 3033–3056.
- [16] R. Sandhu, H.K. Gill, S.K. Sood, Smart monitoring and controlling of pandemic influenza (A/H1N1) using social network analysis and cloud computing, *J. Comput. Sci.* 12 (5) (2015) 11–22.
- [17] J.T. Giger, N.D. Pope, H.B. Vogt, C. Gutierrez, L.A. Newland, J. Lemke, M.J. Lawler, Remote patient monitoring acceptance trends among older adults residing in a frontier state, *Comput. Hum. Behav.* 44 (2) (2015) 174–182.
- [18] M. Bhatia, S.K. Sood, Temporal informative analysis in smart-ICU monitoring: M-healthcare perspective, *J. Med. Syst.* 40 (8) (2016) 1–15.
- [19] S. Sareen, S.K. Sood, S.K. Gupta, A cloud based seizure alert system for epileptic patients that uses higher-order statistics, *Comput. Sci. Eng.* 18 (5) (2016) 56–67.
- [20] M.S. Hossain, G. Muhammad, Cloud assisted industrial internet of things (IIoT) – enabled framework for health monitoring, *Comput. Netw.* 101 (3) (2016) 192–202.
- [21] T.N. Gia, M. Jiang, A.M. Rahmani, T. Westerlund, P. Liljeborg, H. Tenhunen, Fog computing in healthcare internet of things: a case study on ECG feature extraction, 2015 IEEE International Conference on Computer and Information Technology (2015), doi:<https://doi.org/10.1109/CIT/IUCC/DASC/PICOM.2015.51>.
- [22] C.S. Nandyala, H. Kim, From cloud to fog and IoT-based real-time U-healthcare monitoring for smart homes and hospitals, *Int. J. Smart Home* 10 (2) (2016) 187–196.
- [23] MathWorks. Available from: <http://www.mathworks.in/help/fuzzy/examples/fuzzy-c-means-clustering.html> (accessed 06.04.17).
- [24] J. Scott, Social Network Analysis, vol. 19, Sage Publications, 2012.
- [25] The DH Program Demographic and Health Surveys, Survey Dataset Files, (2017) Available from: http://www.dhsprogram.com/data/dataset/Tajikistan_StandardDHS_2012.cfm?flag=1 (accessed 25.04.17).
- [26] The DH Program Demographic and Health Surveys, Survey Dataset Files, (2017) Available from: http://www.dhsprogram.com/data/dataset/Kazakhstan_StandardDHS_1995.cfm?flag=1 (accessed 25.04.17).

- [27] UN Data: A World of Information, Dry Bulb Temperature, (2017) Available from: <http://data.un.org/Data.aspx?d=CLINO&f=ElementCode%3a01> (accessed 25.04.17).
- [28] UN Data: A World of Information, Relative Humidity, (2017) Available from: <http://data.un.org/Data.aspx?d=CLINO&f=ElementCode%3a11> (accessed 25.04.17).
- [29] UN Data: A World of Information, Rainfall, (2017) Available from: <http://data.un.org/Data.aspx?d=CLINO&f=ElementCode%3a39> (accessed 25.04.17).
- [30] S. Tisue, U. Wilensky, NetLogo: design and implementation of a multi-agent modeling environment, *Proc. Agent* (2004) 7–9.