



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Smart monitoring and controlling of Pandemic Influenza A (H1N1) using Social Network Analysis and cloud computing

Rajinder Sandhu*, Harsuminder K. Gill, Sandeep K. Sood

Department of Computer Science and Engineering, Guru Nanak Dev University, Regional Campus, Gurdaspur, Punjab, India



ARTICLE INFO

Article history:

Received 12 April 2015

Received in revised form 30 October 2015

Accepted 4 November 2015

Available online 10 November 2015

Keywords:

Pandemic Influenza A

H1N1

Bioinformatics

Cloud computing

Swine flu

Social Network Analysis

ABSTRACT

H1N1 is an infectious virus which, when spread affects a large volume of the population. It is an airborne disease that spreads easily and has a high death rate. Development of healthcare support systems using cloud computing is emerging as an effective solution with the benefits of better quality of service, reduced costs and flexibility. In this paper, an effective cloud computing architecture is proposed which predicts H1N1 infected patients and provides preventions to control infection rate. It consists of four processing components along with secure cloud storage medical database. The random decision tree is used to initially assess the infection in any patient depending on his/her symptoms. Social Network Analysis (SNA) is used to present the state of the outbreak. The proposed architecture is tested on synthetic data generated for two million users. The system provided 94% accuracy for the classification and around 81% of the resource utilization on Amazon EC2 cloud. The key point of the paper is the use of SNA graphs to calculate role of an infected user in spreading the outbreak known as Outbreak Role Index (ORI). It will help government agencies and healthcare departments to present, analyze and prevent outbreak effectively.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

According to the World Health Organization (WHO), United States (US) spends world's highest percentage of gross domestic budget for health related issues on their citizens [1]. Healthy citizens contribute to nation's development in large extent. Every government has dedicated big amount of budgets for Information Technology (IT) based healthcare support services. IT can provide efficient methods to store and convey information related to the health of every citizen. So, sincere efforts are required to provide ubiquitous healthcare support services using IT infrastructure. Major challenge in any nationwide healthcare support services is the requirement of huge amounts of storage and fast computational power along with its day-to-day maintenance issues. However, with the involvement of distributed technologies such as cloud computing, it is very easy to deploy and use IT systems with varying computational demands. Cloud computing is a highly cost effective IT paradigm which provides on-demand IT resources on a pay-as-use pricing model. Successful adoption of cloud computing by the reputed organizations has inspired researchers to provide healthcare support services over the cloud computing infrastructure [2].

Leveraging cloud computing for healthcare support services facilitates many additional benefits such as ease in swapping records between hospitals and local clinics as well as in effective management of centralized medical records. It not only relieves user from storage and processing infrastructure capital investment, but also prevents maintenance and licensing issues [3]. Accessibility, scalability, availability, storage capability, cost effectiveness and agility are some of the major benefits that can encourage any nation's healthcare departments or private hospitals to shift towards e-health clouds.

The most crucial part of any healthcare support service is the continuous monitoring of an outbreak for any deadly disease. H1N1 is an infectious virus which on spreading affects a large volume of the population. This infectious disease had affected global population in the past due to which approximately 5% of the total world's population had died and hence it had proved to be the deadliest outbreak ever. Alone in the 20th century, the H1N1 virus outbreak has occurred thrice which had affected 500 million people all over the world [1]. WHO states non-existence of primary H1N1 immunity in humans is the root cause of its epidemic. People need to get educated about precautions and handling of this deadly disease. There is a need of using both pharmaceutical as well as non-pharmaceutical strategies to decrease its harmful effects. Cloud computing technologies can be efficiently employed to prevent and keep track of this virus strain. Cloud computing provides virtually

* Corresponding author.

E-mail address: rajsandhu1989@gmail.com (R. Sandhu).

unlimited storage, web services and processing capabilities so, huge amount of virus related data can be linked through social networking and other health aided cloud web services. Moreover, using cloud computing for healthcare support services enables an individual to manage their own health status more effectively. Vital statistics of whole country regarding the epidemic at different level of abstraction can be accessed with high accuracy. Thus, cloud computing can effectively improvised the detection of H1N1 virus at its initial stages. The objectives of the study conducted in this paper is to provide feedback about the individual health condition, design effective information sharing mechanism and real-time representation of the current status of the epidemic. The designed system keeps track of the current status of this pandemic and stores all the records for later use as well. The main contribution of this work is the categorization of patients into different categories, continuous monitoring of infected patients and identifying the probability of each user in spreading or receiving the infection by analyzing his/her relationship with infected patients.

To achieve above said objectives in this paper, a real-time H1N1 monitoring and diagnosis system with large scale data processing and analysis capabilities is proposed using cloud computing. It initially screens H1N1 virus infected patients using random decision tree into five categories based on their respective symptoms. Based on these categories, various control measures are suggested and patients are monitored regularly using a personalized monitoring routine. For every user, Outbreak Role Index (ORI) is calculated which represents his/her future probability in spreading or receiving the infection. Apart from ORI, several other outbreak metrics such as centrality, cluster coefficient, edge embeddeness are also computed using Social Network Analysis. Continuous alerts in case of health threatening conditions are generated and are sent to respective patients. Information and suggestions based upon the respective H1N1 category of user are also provided. The querying capability of this system allows users to intuitively retrieve personalized information. The aim of this paper is to provide a blueprint of the plan to be exercised during this deadly pandemic.

The rest of the paper is organized as follows. Section 2 discuss the related work about H1N1 virus infection and cloud computing in the field of Bioinformatics. Section 3 proposes architecture to monitor, investigate and detect H1N1 virus infection in detail. Section 4 provides experimental results and its discussion. Section 5 concludes the paper.

2. Related work

Related work is divided into three sections which are Pandemic Influenza A (H1N1) virus infection, ICT and mathematical models in H1N1 outbreak and use of cloud computing in the field of Bioinformatics. First section discusses characteristics, causes, and results of H1N1 outbreaks. Second section provides the use of communication and mathematical models in H1N1 infections. Lastly, use of cloud computing in the field of Bioinformatics has been provided.

2.1. Pandemic Influenza A (H1N1) virus infection

H1N1 outbreak in 2009 effected very large number of population across the world. Many authors studied the cause, effect and precaution measures of the 2009 attack. Dawood et al. [4] firstly identified H1N1 influenza virus with about 642 confirmed cases on 15th April 2009. They provided its initial symptoms and effects on the human body. Garten et al. [5] studied antigenic and genetic characteristics of H1N1 virus. Jain et al. [6] examined the clinical characteristics of patients admitted in April to June 2009 affected with H1N1. Neumann et al. [7] reviewed the H1N1 emergence in 2009 and its potential to spread again if not properly controlled.

CDC's Advisory Committee on Immunization Practices (ACIP) [8] also submitted their recommendations on vaccination for H1N1 influenza virus in 2009. As the vaccination of H1N1 is highly recommended for pregnant ladies to prevent infants from infection, McMillan et al. [9] reviewed its need and safety concerns for both mother and infant. Takeda et al. [10] in 2015 studied the effect of latest H1N1 vaccination during pregnancy on both mother and infant. In 2014, WHO [1] piled up health related data for all of its 194 member countries. Progress made in health related targets and associated goals were included. Annual life expectancy and mortality rates due to improper health were estimated by WHO. Data collected by WHO indicated that there is high neonatal mortality rates in low income countries due to improper health guidance. In 2009, Ministry of Health and Family Welfare Pandemic Influenza A (H1N1) [11] also published guidelines on categorization of Influenza A (H1N1) virus infection. Based on various physiological signals, patients were screened and categorized into various categories of infection. According to these categories, relevant information and suggestions were also provided.

2.2. ICT and mathematical models in H1N1 epidemic

Information and Communication Technologies have been extensively used for prevention and early prediction of this deadly epidemic.

Computer science has been also extensively used for analyzing virus epidemiology and its changes. In 2015, Hu [12] examined the change in interdependence of Hemagglutinin and Neuraminidase in H1N1 virus using time series analysis from the data of 2009–2013. In 2014, Kim et al. [13] compared the H1N1, H5N1, H5N2 and H7N9 influenza virus using apriori algorithm. In 2013, Dimitrakopoulou et al. [14] proposed a clustering method known as OLYMPUS for calculating gene expressions. They used data of host response to H1N1 infection for validating the proposed method. In 2013, Bankhead et al. [15] developed a framework in which critical aspects of viral infection are simulated based on spatial temporal data.

Various decision support systems have been proposed by authors for preventing H1N1 outbreaks. In 2015, Lai et al. [16] developed a H1N1 pandemic forecasting system based on daily influenza cases and total population in Hong Kong. They used spatial-temporal and stochastic SEIR methods to predict number of cases for next two days. In 2015, Cruz et al. [17] studied the importance of effective communication system in case of any large scale pandemic. In 2014, Wang et al. [18] examined the spread of H1N1 in 2009 for the schools of China. They found that most of the infection transmission is within the same grade and cross grade infection transmissions are very few. In 2014, Dias and Arruda [19] proposed a mathematical based cost optimization model for controlling H1N1 pandemic based on mean of finite difference method. In 2014, Cui et al. [20] demonstrated the public attention model for epidemics like H1N1 using spatial-temporal patterns. In 2014, Bdeir et al. [21] discovered some of hidden social networks between different organizations during H1N1 epidemics. They also suggested that relevant and timely exchange of information between different organizations can result in better prevention of epidemics. In 2014, Farah et al. [22] developed a dynamic epidemic model for H1N1 using a Bayesian based emulator. Emulator took basic epidemic details and time series of all reported infections to design the model. In 2013, Zhong et al. [23] compared standard as well as modified compartment model of epidemic controlling using H1N1 2009 epidemic data of Arizona state. In 2013, Duan et al. [24] applied a triple approach known as artificial societies, computational experiments and parallel execution to study and control epidemics. They used the data of H1N1 2009 outbreak in China University and examined the social network, student behavior,

population distribution and contact patterns of the virus. In 2012, Yi [25] proposed an ontology driven approach in searching medical databases for case studies in H1N1 outbreaks. In 2011, Bajardi et al. [26] studied the impact of travel restriction during H1N1 2009 epidemic. They found that only restricting international travelers will not result in effective control of epidemic. However, inducing local mobility aware factors can result in reduction of infection spread. In 2009, Balcan et al. [27] modeled the critical demand and supply chain for H1N1 infections antibiotic resources. In 2009, Araz et al. [28] presented a geospatial–temporal disease model to reduce peak infection rates at any instant of a pandemic outbreak. They investigated that if school children are properly scanned and protected, pandemic influenza infections can be minimized with minimum loss of education. In 2009, Coker [29] reviewed swine flu as a communicable disease which spreads by travelling, public gatherings and direct contacts. Virus detection methods such as real time reverse transcriptase–polymerase chain reaction, viral culture, rapid antigen tests, immunofluorescent antibody testing were discussed. Case definition, preventions, recommendations, and efforts at global level were also provided. In 2009, Colizza et al. [30] estimated the H1N1 cases in Mexico City based on a spatially structured epidemic model. In 2009, Sloot et al. [31] developed a decision support system for coordinating any viral infection spread. It provides different type of functionalities such as personalized drug ranking, spatial, temporal, and social interactions.

Several authors focused on determining the flow of infection and simulating the outbreak using different network tools. In 2010, Balcan et al. [32] developed a spreading model that can be used for different epidemics based on spatial temporal data. Gonzalez-Parra et al. in 2011 [33], 2012 [34], and 2015 [35] used spatial–temporal data, fractional order model and complex networks in distributed environment respectively for simulating any H1N1 infection outbreak. In 2011, Van den Broeck [36] studied an effective epidemic modeling and computational tool known as GLEaMViz. It provides different simulating environments as well as visualization tools to effectively represent any epidemic causes and effects. In 2012, Li et al. [37] used data driven inference algorithms to simulate the H1N1 infection outbreak. In 2012, Cheng et al. [38] demonstrated H1N1 outbreak simulations using agent based artificial transport method. In 2012, Tizzoni et al. [39] argued that mathematical based outbreak simulation's reliability is under debate. They tested reliability of Monte Carlo maximum likelihood analysis based mathematical approach using 2009 H1N1 epidemic. Results indicated that simulations were successful in predicting the outbreak spread when compared with real life data. In 2012, Zarrabi et al. [40] proposed a filter-reduction method to study the HIV infection using social interactions and genetic information. In 2014, Duan et al. [41] studied the spreading velocity of any epidemic based on topology of weighted graph created. In 2014, Mei et al. [42] studied the effect of individual behavior during any epidemic using fuzzy cognitive map study. They simulated their framework for H1N1 infection and found that infection can be controlled at low level based on decision making of individuals.

2.3. Cloud computing in the field of bioinformatics

Cloud computing has useful applications in healthcare and its support services due to its capability of handling varying demands. In 2011, Tsai et al. [43] designed a cloud based virtual screening services known as IScreen, for scanning patients and drug designs in Taiwan. In 2012, Zhou et al. [44] reviewed the benefits and characteristics of using cloud computing in bioinformatics. They discussed different service models of cloud computing that can help in bioinformatics field of study. In 2012, Lestrai et al. [45] developed a cloud based software for early diagnosis of lung diseases. They tested it on influenza and tuberculosis patients and results stated its high

reliability. In 2014, Shanahan et al. [46] proposed a framework which uses R statistical tool on Microsoft Azure cloud for bioinformatics data analysis. In 2015, Kwon et al. used cloud computing for analysis of next generation sequencing. In 2015, Griebel et al. [47] examined the use of cloud computing in healthcare sector using multiple research articles. These publications were divided in to six separate categories and results are formatted for each category separately. In 2015, Sandhu et al. [48] developed a cloud based system for predicting and preventing MERS-CoV outbreak using GPS based re-routing.

In 2014, Abbas and Khan [3] stated cloud computing as a prominent paradigm in healthcare services in addition to other business domains. Cloud computing concatenated with healthcare and its support services reduces management as well as maintenance costs. Certain issues such as continuous availability of e-health data and infrastructure to be used for e-health clouds such as public, private or hybrid were also highlighted. In 2014, Sood and Sandhu [49] proposed a proactive resource provisioning methodology using two-dimensional resource provisioning matrix. Resources required in the future were predicted using these matrices by the effective use of artificial neural network. Bill calculated by the cloud providers was cross-checked by independent authorities making this model more cost effective and reliable for mobile cloud users. This approach can be useful in handling mobile data traffic of pandemic outbreak entries. In 2014, Wang et al. [2] proposed a hybrid mobile cloud computing model for medical monitoring, efficient diagnosis, large scale data processing, energy efficiency and execution efficiency. They argued that using their model, patients can manage their own health status in most convenient and affordable manner. In 2014, Xu et al. [50] furnished a semantic data model in order to store and access ubiquitous Internet of Things (IoT) data. They also developed a system that supports emergency medical services and provides flexibility to the user. This method is useful in a distributed heterogeneous data environment and is beneficial for healthcare support service providers to access large ranges of data sources in the mobile application environment. In 2015, Elsebakshi et al. [51] proposed a function network based optimization method for large scale medical data.

As represented in the related work, cloud computing can be considered as an influential paradigm to be collaborated with healthcare support services. More ubiquitous data services for efficient monitoring and diagnosis of infections during a pandemic can be efficiently provided using cloud computing. There is a need for a system that can detect the infection in the early stages and provides continuous monitoring during outbreak. Lack of information and availability of proper treatment at the time of the outbreak is a basic challenge in traditional healthcare support systems. Next section proposes architecture for effectively controlling this deadly epidemic by seamless integration of data mining, cloud computing and social network analysis.

3. Proposed architecture

Fig. 1 shows proposed architecture to effectively detect and monitor H1N1 virus. The data acquisition component is responsible for collecting data of multiple H1N1 and personal attributes. Data can be collected from various sources such as hospitals, individual users, RSS feeds, activity maps and government aided healthcare services. It can also be uploaded to the cloud using web browsers or mobile application. Data is stored in a cloud based repository from where retrieval of data for processing and effective decision making can be efficiently made. Community cloud is preferred for the proposed architecture so that information can be shared securely and effectively among similar groups of agencies who are involved in detecting and controlling the outbreak. Personal details of users are

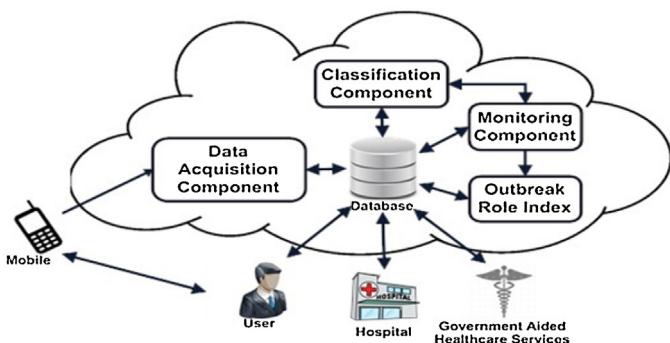


Fig. 1. Proposed architecture to detect and monitor H1N1 virus.

Table 1
Personal details of an individual.

S. no.	Details	Data type
1.	Name	String
2.	Age	Integer
3.	Gender	String
4.	Mobile number	Integer
5.	Family's mobile number (multiple)	Integer
6.	Address	String

encrypted whereas all other information regarding virus attributes is public. After collecting all the relevant information about the user and parameters related to H1N1 virus, users are classified as infected or uninfected using the classification component. Infected patients are categorized into category A, category B (i), category B (ii) and category C as standard set by WHO [11]. According to the classified infected category of user, patients are continuously monitored using monitoring component. Patients with a severe category of infection are monitored more frequently than patients with less severe category of infection. An Outbreak Role Index (ORI) is calculated to find the probability of an individual to spread or receive the infection. Different outbreak metrics are computed using SNA approach. These metrics help government agencies to access situation more effectively. Information and suggestion box stores and communicate information about nearest hospitals and government initiated swine flu camps, first aid and alert in case of health threatening conditions using email, messages, RSS feeds, etc. All component's algorithm and procedures are discussed in detail ahead.

3.1. Data acquisition component

The data acquisition component is responsible for collecting individual as well as generic data about the outbreak. Each user uploads data related to him/her only. But, generic data from RSS

feeds, awareness messages, and flu activity maps can also be collected to further enhance the system capability. Doctors from various hospitals can also upload the information about their patients, their treatment progress and other latest information regarding the H1N1 virus. Whereas doctors are not allowed to upload any personal information of the user such as his/her home address, ethnicity. Data can also be uploaded by government initiated healthcare and support services. Following is the information acquired by system of each user. **Table 1** shows the personal attributes of an individual which has general information about the user. Mobile number of a user will be used to generate a unique patient identification number which will be only used in all future communications.

Table 2 shows H1N1 related attributes and their respective answers collected from every user. These attributes are classified as primary, secondary, tertiary and High Risk Condition (HRC) symptoms. Primary symptoms are the major symptoms which indicate presence of H1N1 virus infection in any patient. Secondary symptoms may be present in any patient as a consequence of primary symptoms. Tertiary symptoms are red flag signs which demonstrate more worsening conditions for a virus infected patient so that immediate hospitalization and treatment should be instructed. A person with HRC symptoms indicates that he/she can be more prone towards H1N1 virus infection. A person with positive HRC symptoms can be affected more vigorously as compared to persons who do not possess any HRC symptoms.

3.2. Classification component

This component classifies the user depending on H1N1 attribute data using random decision tree as uninfected category, category A, category B (i), category B (ii) or category C. Random decision tree was generated using Weka 3.7 [52].

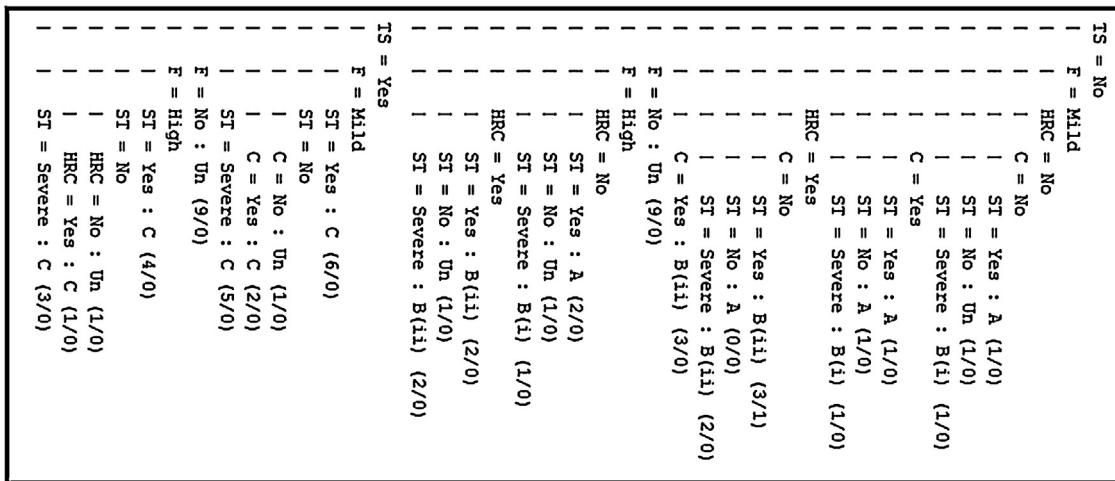
As shown in Fig. 2 generated by Weka 3.7, the user is categorized as infected under category A if he/she has mild fever along with cough or sore throat. Category A infected patients may also show secondary symptoms such as headache, body ache. However, if fever and sore throat are very severe, then the category of infection will be B (i). Category B (ii) includes infected patients showing HRC symptoms along with symptoms of category B (i). Category C of infection shows tertiary symptoms along with all the symptoms covered under category A, category B (i) and category B (ii). The user is uninfected if he/she does not satisfy above-mentioned conditions. Categorization is done under the guidance of WHO [11].

3.3. Continuous monitoring of H1N1 virus infected users

Continuous monitoring refers to regularly examine the treatment and symptoms of individual users, so that appropriate progress report of each patient can be maintained in the system. According to infected categories of patient as classified by random

Table 2
Symptoms of H1N1 virus.

H1N1 virus symptoms							
Primary symptoms		Secondary symptoms		Tertiary symptoms		High risk condition symptoms	
Fever	High/Mild/No	Diarrhea	Yes/No	Breathlessness	Yes/No	Lung disease	Yes/No
Cough	Yes/No	Headache	Yes/No	Drowsiness	Yes/No	Heart disease	Yes/No
Sore throat	Yes/No	Body ache	Yes/No	Chest pain	Yes/No	Liver disease	Yes/No
		Vomiting	Yes/No	Sputum mixed in blood	Yes/No	Neurological disorder	Yes/No
				Fall in blood pressure	Yes/No	Kidney disorder	Yes/No
				Bluish discoloration of nails	Yes/No	Blood disorder	Yes/No
						Diabetes	Yes/No
						HIV/Aids	Yes/No
						Cancer	Yes/No
						Long term cortisone therapy	Yes/No
						Chronic disease	Yes/No



TS- Tertiary Symptoms, F- Fever, HRC- High Risk Conditions, C- Cough, ST- Sore Throat, A- Category A, B(i)- Category B(i), B(ii)- Category B(ii), C- Category C, Un- Uninfected.

Fig. 2. Categorization using decision tree.

decision tree, monitoring is done for different period of times. **Table 3** shows monitoring interval of infected patients. However, monitoring interval can also be changed by consulting a specialized doctor.

The process of continuous monitoring and examining is repeated until the patient is categorized as uninfected. Continuous alerts and notifications generated for an infected user are sent on his/her mobile phones. In case any user is under the age of 18, notifications are also sent to his/her registered guardian. Alerts are also sent to nearest hospitals and government aided health care services about the infected patient's condition and progress report.

Progress report of the patient is also uploaded regularly by doctors and healthcare agencies under the unique Patient Identity Number (PIN) provided by the system. User can also ask a query or enter symptoms after regular intervals. Decision tree will again predict category of the patient referred as Outcome of Decision Tree (ODT) as shown in Algorithm 1. If there is no change in previous and current category of the infection, same routine will be followed, but if the category is changed, i.e. either improved or worsen then database corresponding to that patient is immediately updated, sudden checkup is required to change the treatment and monitoring process is also updated. Alerts are sent according to the new category of infection to individual and respective doctor.

Algorithm 1. Rechecking symptoms and category of the patient

```

Input: PIN and symptoms;
Output: Revised category of infection.
begin
  login using PIN;
  Record symptoms as per desired date;
  Use decision tree to predict new category based on updated symptoms;
  if new category=previous category do
    provide next monitoring date;
    save database;
  else
    update monitoring interval;
    update category of the patient;
    alert doctor and fix an appointment;
    provide next monitoring date to patient;
  end if

```

3.4. Outbreak prevention

In the previous sections, proposed architecture worked at the individual level. In this section, spreading of infection and its prevention measures are addressed at a higher level of abstraction.

Outbreak prevention is an effort to represent H1N1 outbreak, current state and flow so that it can be stopped from being spread further.

A Social Network Analysis (SNA) graph is proposed which represents each individual as a node and edges are formed between users who are in contact with each other. Using this technique, the infected patients and their connections with other infected or uninfected user can be depicted effectively. SNA based outbreak prevention and its different metrics are explained in detail in Section 4.

3.5. Information and suggestion box

Information and suggestion box is a component which is used to deliver the latest information and suggestions about H1N1 virus infection. Information and suggestions can be sent to individual as SMS or e-mail address. **Tables 4 and 5** show various suggestions for uninfected and infected users respectively. Suggestions are provided according to the guidelines provided by WHO [11]. Various governments aided healthcare services, indulged in prevention of the outbreak can suggest latest information regarding pandemic, latest treatment methodologies and highly affected geographical area, etc. Comments and questions can also be queried by the user from this component which can be used to conduct complete user satisfaction survey.

4. SNA based H1N1 outbreak prevention

SNA uses many network techniques to analyze large and complex social networks for example Facebook or political views of citizens [53]. SNA provides various tools and techniques to analyze and extract results from complex graphs. In the proposed architecture, identifying H1N1 patients is important, but more important is to prevent the spreading of the H1N1 outbreak. A clear indication of H1N1 outbreak range and patients interlinking will help

Table 3
Monitoring interval of infected patients.

Category	Monitoring interval
Category A	24–48 h
Category B (i)	12 h
Category B (ii)	10 h
Category C	6 h

Table 4

Suggestions for uninfected users.

S. no.	Suggestions
1.	Avoid contact with infected people
2.	Take vaccines and antiviral drugs, they do not cure the flu, but help you to get rid of morbidity and mortality
3.	Four antiviral drugs are approved, oseltamivir, zanamivir, amantadine, rimantadine
4.	If you are developing symptoms indicating flu, isolation must be provided from other people of the family, public transport and work places
5.	Cover mouth and nose while sneeze and cough with tissue and bin the tissue after one use
6.	Wash hands frequently with soap and hot water or use alcohol-based hand sanitizer
7.	Keep yourself up to date with the latest information on television, internet, newspapers and radios that how fast flu is spreading and how you can protect yourself and your family from the flu, what precautions and treatment to be taken.

- = Uninfected Category (Dark Green)
- = Category A (Light Green)
- = Category B(i) (Orange)
- = Category B(ii) (Brown)
- = Category C (Red)

Fig. 3. Coloring scheme for representing different categories in the global SNA graph.

government and healthcare agencies to stop the outbreak at the lowest possible level. It will also provide particular areas and individuals to focus for government healthcare agencies.

4.1. Creating global SNA graph

Tools and techniques provided by SNA are useful if a social network graph has been created efficiently. Social network graph created for representing H1N1 outbreak in proposed architecture is known as a global SNA graph. Once a user is screened as infected by the classification component, he/she and his/her mobile list contacts are added to the global SNA graph. Algorithm 2 provides step-by-step computations for creating an effective global SNA graph.

Using Algorithm 2, a global SNA graph will be created and regularly updated with time. Many important conclusions can be made from this graph which will help government health agencies to control the H1N1 outbreak. Fig. 3 shows the coloring scheme of nodes used for different categories, while creating global SNA graphs.

Table 5

Suggestions for infected users.

S. no.	Category A infection	Category B infection	Category C infection
1.	Patient should be treated only for symptoms and no other treatment is required.	Oseletamivir treatment should be given.	Immediate consultation of doctor, treatment and hospitalization is required.
2.	No test for the confirmation of H1N1 virus is required.	Confirmation test for H1N1 virus is not required.	
3.	Isolation is required from uninfected and HRC members.	Isolation is required from uninfected and HRC members.	
4.	Patient should be reconsidered by a doctor after 24–48 h.		

Algorithm 2. Creating global SNA graph

```

Input: Infected User and Mobile.Contact[N];
Output: New or updated global SNA graph
Begin
Get the category of user;
If user is categories as infected do
Identify the category of infection;
Create a new node at the residence address of the user with color same as his/her category;
Scan mobile contact list of the user;
for i=0 to N – 1 do
if Mobile.Contact[i] is already in the graph do
Create a new edge between Mobile.Contact[i] and user;
Else
Create a new node with Mobile.Contact[i];
Create edge between Mobile.Contact[i] and user;
end if
end for
end if

```

4.2. Outbreak Role Index

In the proposed architecture, an index known as Outbreak Role Index (ORI) is proposed which calculates the probability of any user to receive or spread the infection. ORI is calculated from the global SNA graph. ORI states the ratio of the number of infected people in contact with user to the total number of mobile contact list users. ORI of ith user can be stated as

$$\text{ORI}_i = \frac{\text{#infected people in contact with patient}}{\text{Total number of people in contact with patient}}$$

$$\text{ORI}_i = \frac{N - \text{Edges with Green Nodes}}{N}$$

where N is total number of the edges of ith node.

Two types of user which will require special attention of healthcare agencies are:

- If the user is infected and his/her ORI is low, it means people in contact with this particular user are not infected so this user posses high probability of spreading the infection. These types of users should be quarantined as early as possible. Further, this type of user is represented as Spreading User (SU).
- If the user is uninfected and his/her ORI is high. It means persons in contact with this particular user are almost infected with H1N1. This posse's very high risk of particular user to get infected. Appropriate risk measures and alerts should be sent to these types of users. Further, this type of user is referred as Getting User (GU).

4.3. Outbreak metrics

Some of the important definitions and outputs that can be drawn from the global SNA graph are discussed in this section.

Centrality of any node in the global SNA graph is very effective metric to represent the level of involvement of any user in

spreading the outbreak. High centralized infected users will have a high probability of spreading the outbreak.

Definition 1. Centrality of any user

Centrality will represent the number of connections each user has on the global SNA graph. For any global SNA graph $G = (V, E)$, let x^* user has the highest centrality. Then centrality of graph G can be computed as

$$C(G) = \frac{\sum_{i=1}^V [C(x^*) - C(x_i)]}{(N-1)(N-2)}$$

where N is the total number of users.

Definition 2. Closeness with infected user

Sometimes the user is not in direct contact with the H1N1 infected user, but knows another user who is in direct contact with H1N1 infected users. The closeness will study this relationship and will derive any user's shortest distance from infected and uninfected users. Closeness of any i th user with infected and uninfected patients can be calculated as

$$\text{Closeness}_i(\text{Infected}) = \frac{\sum_{t \in V/u} d_G(i, t)}{N-1}$$

and

$$\text{Closeness}_i(\text{Uninfected}) = \frac{\sum_{t \in u} d_G(i, t)}{N-1}$$

where $d_G(i, t)$ is the shortest distance between user i and t , u are the uninfected users, V is the set of all users and N is the total number of users in the global SNA graph.

Definition 3. Relative score of each user

The relative score of a user will help to identify critical users from the global SNA graph. Appropriate alerts can be generated based on this score value. Relative score S_i of any user i can be computed as

$$S_i = \frac{1}{\lambda} \sum_{j \in V(i)} S_j = \frac{1}{\lambda} \sum_{j=1}^N A_{ij} * S_j$$

where λ is the eigen constant, A_{ij} is the adjacency matrix and $V(i)$ is the set of nodes which are connected to i th node. A user with high score should be quarantined as early as possible.

Definition 4. Chance of forming clusters

Finding the probability for the formation of clusters of H1N1 infected patients in any region is of high importance. It will help the government agencies to seize that area and eliminate all kinds of travel from that region. Cluster Coefficient (CC) value is used in the proposed architecture to identify the probability of cluster formation. Global value of CC can be computed as

$$CC(\text{Global}) = \frac{3 * \# \text{triangles}}{\# \text{Connected Tripets}}$$

CC can be calculated for the individual user as follows:

$$CC_i = \frac{\# \text{connection between neighbor users}}{N_i * (N_i - 1)/2}$$

where N_i the number of neighbor node is i th user has. CC of whole graph can be computed by the summation of all individual CC as follows:

$$CC(\text{Whole Graph}) = \frac{1}{N} \sum_i CC_i$$

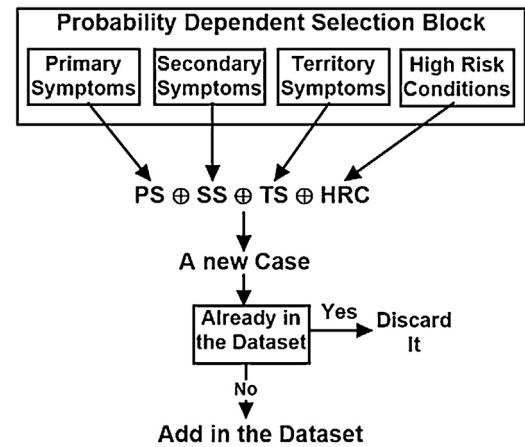


Fig. 4. Steps involved in creation of synthetic H1N1 dataset based on symptoms.

where N is the total number of users. Edge embeddeness represents the level of connection between neighbor of any user. Edge embeddeness can be computed as follows:

$$EE_{ij} = \frac{\# \text{user neighbor of } i \text{ and } j}{\# \text{user neighbor of atleast } i \text{ or } j}$$

Greater the value of EE_{ij} higher the chance of the cluster formation so, high probability of the spreading of the H1N1 outbreak.

5. Experiment setup and performance analysis

After depth search on the internet and email correspondence with some of the doctors involved with H1N1 flu, we were not able to receive any H1N1 flu infection database based on symptoms of patients to test our proposed architecture. For experimental setup and performance analysis of the proposed architecture, synthetic data are generated in conjunction with Dr. Pankaj Sood. To test the proposed architecture effectively, experimental section is divided into four main components which are

- Creation of synthetic data for H1N1.
- Training and testing of synthetic data using random decision tree.
- Performance analysis of proposed model on Amazon EC2 cloud.
- Outbreak metrics using SNA.

5.1. Creation of synthetic H1N1 data

Non-availability of symptoms based dataset for H1N1 flu forced us to create synthetic data for appropriate evaluation of the proposed architecture. Synthetic data is generated very systematically so that no possible case has been left out and all possible scenarios are listed. Table 6 shows the probabilities of each H1N1 virus symptoms to be present in any new generated case while creating synthetic dataset for H1N1. Step used to create dataset are listed in Algorithm 3 and graphically explained in Fig. 4. The main objective of probability based case generation is to form all high probability cases based on each symptom's individual probability. Probabilities assigned in Table 6 are used for testing of the proposed architecture with consultation by Dr. Pankaj Sood. However, if required or based on the situation in hand, these values can be changed.

Table 6

Probabilities set for H1N1 virus symptoms to be Yes.

Probability of H1N1 virus symptoms to be Yes						
Primary symptoms		Secondary symptoms		Tertiary symptoms		High risk condition symptoms
Fever	0.16	Diarrhea	0.12	Breathlessness	0.06	Lung disease
Cough	0.17	Headache	0.15	Drowsiness	0.08	Heart disease
Sore throat	0.17	Body ache	0.15	Chest pain	0.05	Liver disease
No symptoms	0.50	Vomiting	0.08	Sputum mixed in blood	0.04	Neurological disorder
		No symptom	0.50	Fall in blood pressure	0.04	Kidney disorder
				Bluish discoloration of nails	0.03	Blood disorder
				No symptoms	0.70	Diabetes
						HIV/Aids
						Cancer
						Long term cortisone therapy
						Chronic disease
						No symptoms
						0.80

Table 7

Detailed Accuracy by Class for Random Decision Tree in Weka 3.7.

TP rate	FP rate	Precision	Recall	F-Measure	MCC	ROC area	Class
1.000	0.000	1.000	1.000	1.000	1.000	1.000	Uninfected
0.938	0.061	0.895	0.923	0.906	0.835	0.885	Category A
0.795	0.204	0.741	0.776	0.769	0.772	0.787	Category B (i)
0.994	0.005	0.921	0.951	0.937	0.872	0.906	Category B (ii)
0.996	0.003	0.940	0.977	0.951	0.900	0.928	Category C
Weighted avg.	0.944	0.054	0.899	0.925	0.912	0.875	0.901

Algorithm 3. Creation of synthetic H1N1 dataset

```

Input: Different H1N1 symptoms, Number of distinct cases required
Output: Dataset
Begin
While (generated cases less than total required cases)
Assign values to primary symptoms based on probability in Table 6;
Assign values to secondary symptoms based on probability in Table 6;
Assign values to tertiary symptoms based on probability in Table 6;
Assign values to high risk condition symptoms based on probability in Table 6;
Create a new case by combining all symptoms values;
if New case is already present do
discard the new case;
else
Add the new case;
Increase the value of generated cases by one;
end if
do

```

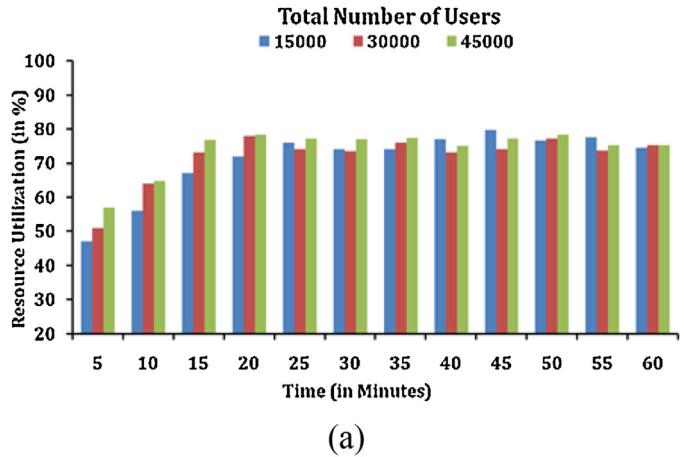
Algorithm 3 has been executed for creating 5000 cases of H1N1 virus. Out of these 5000 cases, high probability 200 cases have been used to train random decision tree in Weka 3.7 [52] on Amazon EC2 cloud instance. Accuracy of created random decision tree is tested on 5000 cases which is explained in detail in Section 5.2.

5.2. Training and testing of decision tree

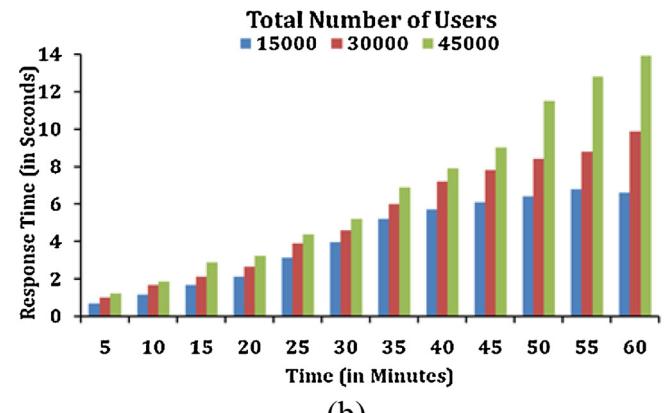
Top 200 high probability cases are used to train the random decision tree in Weka 3.7 with minNum, minVarianceProp and seed equal to 1.0, 0.001 and 1 respectively. Random tree created is shown in Fig. 2.

Various statistical values of the model tested in Weka 3.7 are shown in Tables 7–9. Decision tree classifies the users with an accuracy of 94%. The detailed accuracy of each class parameter as classified by decision tree is also listed in Table 7.

True Positive (TP) rate, also known as sensitivity and False Positive (FP) rate, also known as specificity are statistical measure to rate quality of any classification algorithm. TP rate defines the percentage of categories of H1N1 instances which classification algorithm classifies correctly. FP rate defines the percentage of categories of H1N1 instances which classification algorithm classifies wrongly. A classification algorithm with high TP rate and low FP



(a)



(b)

Fig. 5. Cloud system performance for proposed architecture: (a) resource utilization of system, (b) response time of system.

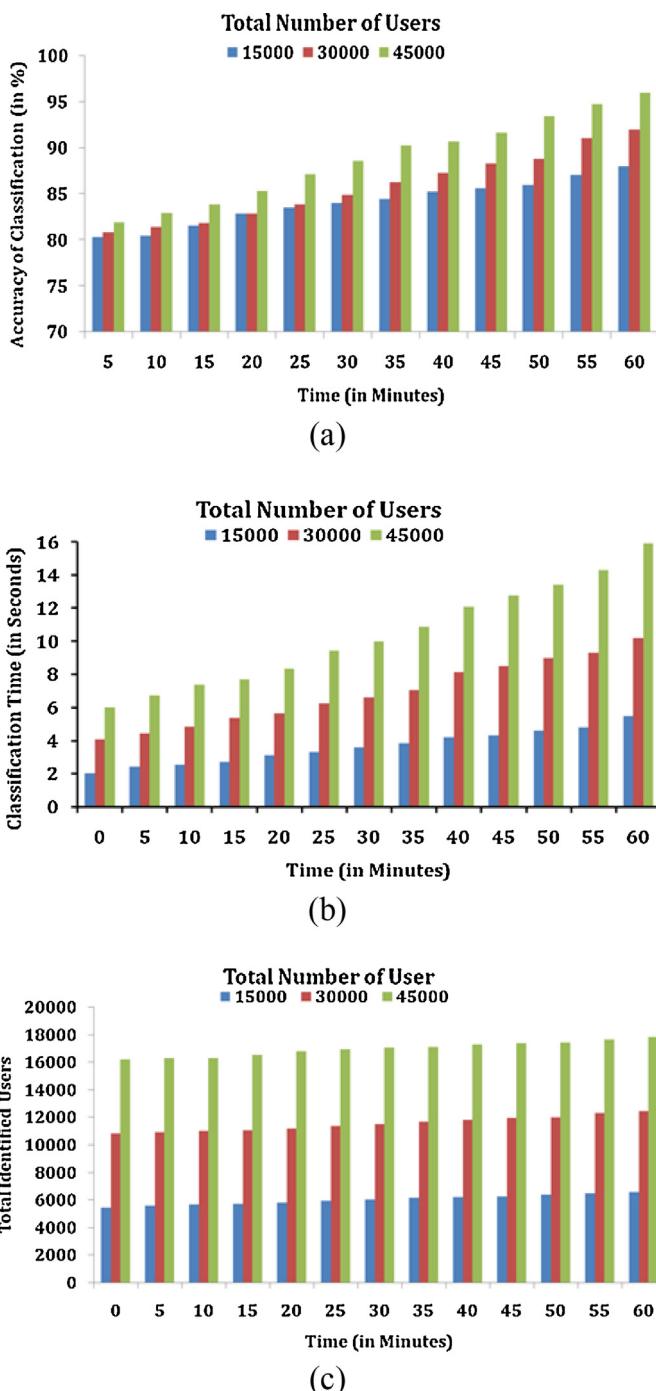


Fig. 6. Performance analysis of classification tool over the Amazon cloud: (a) accuracy of classification, (b) classification time, (c) total infected users classified correctly.

Table 8
Summary of 10-fold cross-validation of random decision tree in Weka 3.7.

Correctly classified instances	47,153 (94%)
Incorrectly classified instances	2847 (6%)
Kappa statistic	0.81
Mean absolute error	0.0621
Root mean squared error	0.2491
Relative absolute error	13%
Root relative squared error	50.3241%
Coverage of cases (0.95 level)	93%
Mean rel. region size (0.95 level)	34.3581%
Total number of instances	50,000

Table 9
Confusion matrix of random decision tree in Weka 3.7.

		Classified category				Actual category
Un	A	B (i)	B (ii)	C		
23,732	0	0	0	0	0	Un
0	11,641	758	0	0	0	A
0	0	8040	2070	0	0	B (i)
0	0	4	3089	13	0	B (ii)
0	0	0	2	650	0	C

rate is desirable. The random decision tree provides a high TP rate that is 0.944 and the low FP rate that is 0.054.

Precision and recall provides the relevancy of the classified instances. Higher the value of precision and recall, better will be classification because only relevant instances will be classified. Proposed algorithm provides very high precision and recall rate values which are 0.899 and 0.925 respectively. F-Measure is a statistical value for classification accuracy which lies between 0 and 1. Algorithm with high F-Measure value will be more accurate and our random decision tree provides F-Measure of 0.912. The high value of ROC area represents the accuracy of the proposed algorithm. With the level of accuracy conceived in above listed statistical terms, it justifies the use of random decision tree for the proposed architecture.

5.3. Performance analysis of Amazon EC2

Application for predicting the category of user depending on his/her symptoms is proposed to host on a cloud so that services are provided without any interruption. User can add the symptoms, change the symptoms, delete the symptoms or search for similar cases in the proposed architecture. The system should provide the desired level of Quality of Service (QoS) to end user while performing these operations. To test this performance of proposed architecture, synthetic generated H1N1 flu cases are stored over shared cloud storage provided by Amazon EC2 [54]. General purpose xlarge [54] clusters are used to set up application over the cloud. It can be accessed by multiple end users such as doctors, users, healthcare departments and government agencies. Maximum number of cases

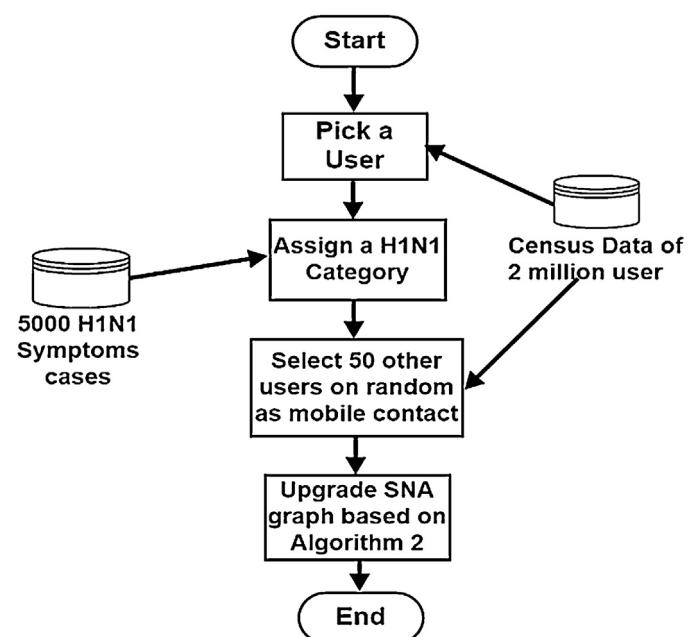


Fig. 7. Generation of 2 millions of H1N1 cases using census data.

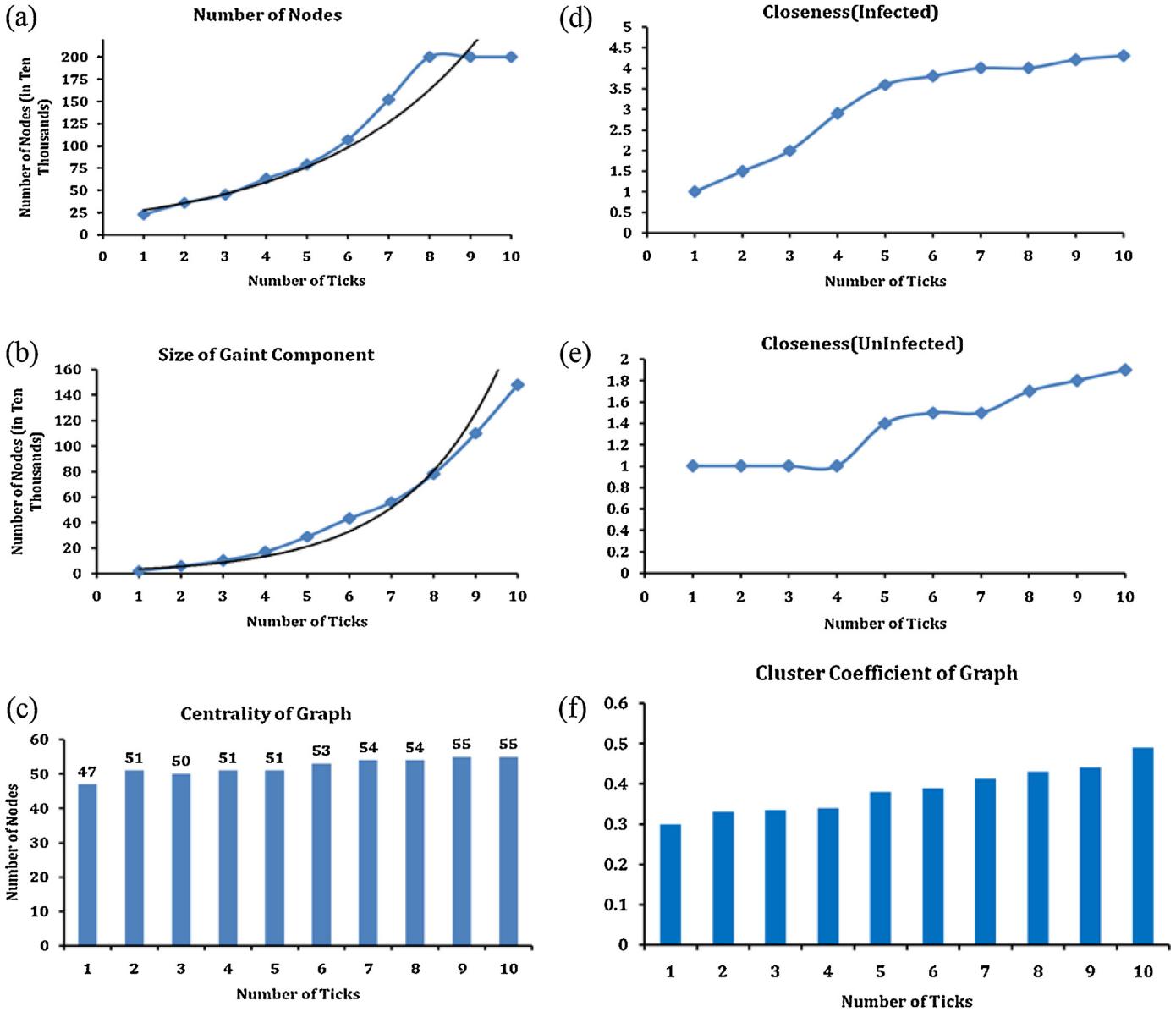


Fig. 8. Different outbreak metrics generated using SNA: (a) total size of global SNA graph, (b) size of giant component of global SNA graph, (c) centrality of global SNA graph, (d) closeness of infected users, (e) closeness of uninfected users, (f) cluster coefficient of global SNA graph.

used to test the performance levels are 50,000. Each data access request is amplified to hundred times by using a JavaScript coded in the application frontend. Total requests to the H1N1 flu database were divided at 40% of search, 35% of update, 15% of insert and 10% of deleting. Initially the system was started with 200 requests, then after each 5-min request to system was increased by 200 and system performance was studied for a total experiment time of 60 min.

Fig. 5(a) shows resource utilization of the proposed architecture for different number of users. The system achieves saturation very fast when numbers of users are 16,000 because larger the database more resources will be required to perform any transaction. Similarly, change in response time is also studies for the system, shown in Fig. 5(b). A system with less number of users provides low response time because there are fewer entries in the database to perform any operation.

Classification algorithm performance is also verified on the Amazon cloud. One of the vital tasks of the proposed architecture

is the classification of users into infected and uninfected, so it requires a high performance level. Some important performance measures related to random decision tree classification algorithm are shown in Fig. 6. Fig. 6(a) shows the accuracy of the classification algorithm. Accuracy of classification algorithm is the percentage of correct classification of uninfected and infected users in their respective category. Random decision tree starts with an accuracy rate of 80%, but as time passes its accuracy increases due to large user dataset. Fig. 6(b) shows the average time took by system for initial analysis. Large user dataset requires more time for initial analysis as depicted in the graph. It is because the system needs more time to compare its prediction with actual H1N1 case result of large user dataset. Fig. 6(c) shows the effectiveness of the system for identifying H1N1 infected patients. Total number of users it predicted as infected are very large as compared to traditional methods. It also shows the total number of lives saved by the system because these users are diagnosed in early stages.

Table 10

Outbreak metrics computed from simulations.

Function	Description
<i>num-nodes</i>	Total number of nodes in global SNA graph, i.e. total number of users to scan.
<i>giant-component-size</i>	Size of highly connected component.
<i>ORI</i>	The Outbreak Role Index is provided for each node and it is kept updated with each tick.
<i>Centrality</i>	The centrality value of whole graph is computed after each tick.
<i>Closeness(<i>infected</i>)</i>	The shortest distance from infected user is calculated if node is itself uninfected with each tick.
<i>Closeness(<i>uninfected</i>)</i>	The shortest distance from uninfected user is calculated if node is itself infected with each tick.
<i>Score</i>	The relative score of each node is also computed with each tick.
<i>Cluster coefficient</i>	The probability of forming a cluster is generated with each tick.
<i>Edge embeddeness</i>	This value is also computed to know the chances of forming the clusters.

5.4. SNA based outbreak prevention

Growth model is simulated in NetLogo 5.6 [55] for simulation of outbreak prevention using SNA techniques. To analyze the system for a large number of agents, synthetic data is generated for 2 million users. Every user acts as agent in NetLogo simulations. H1N1 cases generated in Section 5.1 are now mapped with demographic and geographic data of 2 million of users. Census data of 2 million users [56] is mapped systematically with generated H1N1 cases to create global SNA graph as shown in Fig. 7 (Table 6).

Each time a new user is registered with the system, the system will automatically assign another 50 users as his/her mobile contacts. These 50 mobile contact list users will be chosen randomly from the census data of 2 million users. The generated global SNA graph is evaluated on different parameters. Different parameters used for representing outbreak are listed in Table 9.

Fig. 8 (Table 10) shows different outbreak metrics generated from the global SNA graph created using synthetic dataset. Fig. 8(a) shows the increase in number of nodes in global SNA graph after each tick of NetLogo. Increase in size of the global SNA graph follows an exponential trend as shown by the trend line in Fig. 8(a). A giant component of the global SNA graph follows the trend of number of nodes and follow an exponential trend as depicted in Fig. 8(b). The overall centrality of the graph remains same around 50 in all ticks as shown in Fig. 8(c). This is because linking 50 users from pool of ~300 thousand users will not affect the overall centrality of the global SNA graph. Closeness of infected user is higher than uninfected users as shown in Fig. 8(d) and (e) because system has more uninfected user than infected user. Fig. 8(f) provides the cluster coefficient of the global SNA graph. SNA act as very useful tool to analyze level of outbreak using multiple parameters. Experiment results indicated its suitability in the proposed architecture.

6. Conclusion

Easily transmitted diseases are one of the major concerns of any nation's government and healthcare departments. With the increase in information technologies, it is possible to control many infections in most effective and efficient way. In this paper, architecture is proposed for predicting and preventing an airborne H1N1 flu using cloud computing and Social Network Analysis (SNA). The random decision tree is used to classify users and cloud computing is used for effective information analysis and sharing. The key point of the paper is the use of SNA methods to represent each H1N1 flu infected users on a global SNA graph to show user's dependencies.

Proposed system is implemented on Amazon EC2 cloud which provides 94% accuracy in classification and around 81% cloud resource utilization. It will help uninfected citizens avoid regional exposure and government agencies to address the problem more effectively. Future work will include secure transmission of information among different entities of the proposed architecture. We will also extend the proposed architecture for more generic diseases such as Hypertension, Diabetics and Cancer.

References

- [1] World Health Organisation, World Health Statistics 2014, 2014 [Online]. Available http://apps.who.int/iris/bitstream/10665/112738/1/9789240692671_eng.pdf [accessed 19.07.15].
- [2] X. Wang, Q. Gui, B. Liu, Z. Jin, Y. Chen, Enabling smart personalized healthcare: a hybrid mobile-cloud approach for ECG telemonitoring, IEEE J. Biomed. Health Inf. 18 (3) (2014) 739–745.
- [3] A. Abbas, S.U. Khan, A review on the state-of-the-art privacy-preserving approaches in the e-Health clouds, IEEE J. Biomed. Health Inf. 18 (4) (2014) 1431–1441.
- [4] F.S. Dawood, S. Jain, L. Finelli, M.W. Shaw, S. Lindstrom, R.J. Garten, L.V. Gubareva, X. Xu, C.B. Bridges, T.M. Uyeki, Emergence of a novel swine-origin Influenza A (H1N1) virus in humans, N. Engl. J. Med. 360 (June (25)) (2009) 2605–2615.
- [5] R.J. Garten, C.T. Davis, C.A. Russell, B. Shu, S. Lindstrom, A. Balish, W.M. Sessions, X. Xu, E. Skepner, V. Deyde, M. Okomo-Adhiambo, L. Gubareva, J. Barnes, C.B. Smith, S.L. Emery, M.J. Hillman, P. Rivallier, J. Smagala, M. de Graaf, D.F. Burke, R.A.M. Foucheir, C. Pappas, C.M. Alpuche-Aranda, H. López-Gatell, H. Olivera, I. López, C.A. Myers, D. Faix, P.J. Blair, C. Yu, K.M. Keene, P.D. Dotson, D. Boxrud, A.R. Sambol, S.H. Abid, K. St George, T. Bannerman, A.L. Moore, D.J. Stringer, P. Blevinis, G.J. Demmler-Harrison, M. Ginsberg, P. Kriner, S. Waterman, S. Smole, H.F. Guevara, E.A. Belongia, P.A. Clark, S.T. Beatrice, R. Donis, J. Katz, L. Finelli, C.B. Bridges, M. Shaw, D.B. Jernigan, T.M. Uyeki, D.J. Smith, A.I. Klimov, N.J. Cox, Antigenic and genetic characteristics of swine-origin 2009 A (H1N1) influenza viruses circulating in humans, Science 325 (July (5937)) (2009) 197–201.
- [6] S. Jain, L. Kamimoto, A.M. Bramley, A.M. Schmitz, S.R. Benoit, J. Louie, D.E. Sugerman, J.K. Druckenmiller, K.A. Ritger, R. Chugh, S. Jasuja, M. Deutscher, S. Chen, J.D. Walker, J.S. Duchin, S. Lett, S.E. Soliva, V. Wells, D. Swerdlow, T.M. Uyeki, A.E. Fiore, S.J. Olsen, A.M. Fry, C.B. Bridges, L. Finelli, Hospitalized patients with 2009 H1N1 influenza in the United States, April–June 2009, N. Engl. J. Med. 361 (November (20)) (2009) 1935–1944.
- [7] G. Neumann, T. Noda, Y. Kawaoka, Emergence and pandemic potential of swine-origin H1N1 influenza virus, Nature 459 (June (7249)) (2009) 931–939.
- [8] Use of influenza A (H1N1) 2009 monovalent vaccine: recommendations of the Advisory Committee on Immunization Practices (ACIP), Morb. Mortal. Wkly. Rep. 58 (RR-10) (2009) 1–10.
- [9] M. McMillan, D. Kralik, K. Porritt, H. Marshall, Influenza vaccination during pregnancy: a systematic review of effectiveness and adverse events no, Database Syst. Rev. Implement. Rep. 12 (6) (2014 Jul) 251–381.
- [10] S. Takeda, M. Hisano, J. Komano, H. Yamamoto, H. Sago, K. Yamaguchi, Influenza vaccination during pregnancy and its usefulness to mothers and their young infants, J. Infect. Chemother. 21 (4) (2015 Apr) 238–246.
- [11] G. MOHFW, Guidelines on Categorization of Influenza A H1N1 Cases During Screening for Home Isolation, Testing, Treatment and Hospitalization, 2009.
- [12] W. Hu, Computational study of interdependence between hemagglutinin and neuraminidase of pandemic 2009 H1N1, IEEE Trans. Nanobiosci. 14 (March (2)) (2015) 157–166.
- [13] D.Y. Kim, H.J. Kim, J. Bae, T. Yoon, Trends and Applications in Knowledge Discovery and Data Mining, vol. 8643, Springer International Publishing, Cham, 2014.
- [14] K. Dimitrakopoulou, A.G. Vrahatis, E. Wilk, A.K. Tsakalidis, A. Bezerianos, OLYMPUS: an automated hybrid clustering method in time series gene expression. Case study: host response after Influenza A (H1N1) infection, Comput. Methods Programs Biomed. 111 (September (3)) (2013) 650–661.
- [15] A. Bankhead, E. Mancini, A.C. Sims, R.S. Baric, S. McWeeney, P.M.A. Sloot, A simulation framework to investigate *in vitro* viral infection dynamics, J. Comput. Sci. 4 (May (3)) (2013) 127–134.
- [16] P.-C. Lai, C.B. Chow, H.T. Wong, K.H. Kwong, Y.W. Kwan, S.H. Liu, W.K. Tong, W.K. Cheung, W.L. Wong, An early warning system for detecting H1N1 disease outbreak – a spatio-temporal approach, Int. J. Geogr. Inf. Sci. (May) (2015) 1–18.
- [17] M.A. Cruz, N.M. Hawk, C. Poulet, J. Rovira, E.N. Rouse, Public health incident management: logistical and operational aspects of the 2009 initial outbreak of H1N1 influenza in Mexico, J. Emerg. Manag. (Weston, MA) 13 (January (1)) (2015) 71–77.
- [18] L. Wang, C. Chu, G. Yang, R. Hao, Z. Li, Z. Cao, S. Qiu, P. Li, Z. Wu, Z. Yuan, Y. Xu, D. Zeng, Y. Wang, H. Song, Transmission characteristics of different students during a school outbreak of (H1N1) pdm09 influenza in China, 2009, Sci. Rep. 4 (January) (2014).
- [19] C.M. Dias, E.F. De Arruda, Computational cost optimization for influenza A (H1N1) epidemic model, in: Proceedings of 7th International Congress on

- Environmental Modelling and Software: Bold Visions for Environmental Modeling**, iEMSS, vol. 3, 2014, pp. 1319–1325.
- [20] K. Cui, X. Zheng, Z. Zhang, D. Zeng, *Smart Health*, vol. 8549, Springer International Publishing, Cham, 2014.
- [21] F. Bdeir, L. Hossain, J. Crawford, S. Carlsson, Inter-organisational coordination of H1N1 outbreak: data collection, and analyses of a pilot field study, *J. Decis. Syst.* 23 (April (2)) (2014) 151–166.
- [22] M. Farah, P. Birrell, S. Conti, D. De Angelis, Bayesian emulation and calibration of a dynamic epidemic model for A/H1N1 Influenza, *J. Am. Stat. Assoc.* 109 (December (508)) (2014) 1398–1411.
- [23] W. Zhong, Y. Kim, M. Juhn, Modeling dynamics of an influenza pandemic with heterogeneous coping behaviors: case study of a 2009 H1N1 outbreak in Arizona, *Comput. Math. Organ. Theory* 19 (November (4)) (2012) 622–645.
- [24] W. Duan, Z. Cao, Y. Wang, B. Zhu, D. Zeng, F.-Y. Wang, X. Qiu, H. Song, Y. Wang, An ACP approach to public health emergency management: using a campus outbreak of H1N1 influenza as a case study, *IEEE Trans. Syst. Man. Cybern. Syst.* 43 (September (5)) (2013) 1028–1041.
- [25] M. Yi, Effective medical resources searching using an ontology driven medical information retrieval system, *Electron. Libr.* 30 (April (2)) (2012) 248–263.
- [26] P. Bajardi, C. Poletto, J.J. Ramasco, M. Tizzoni, V. Colizza, A. Vespignani, Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic, *PLoS ONE* 6 (January (1)) (2011) e16591.
- [27] D. Balcan, V. Colizza, A.C. Singer, C. Chouaid, H. Hu, B. Gonçalves, P. Bajardi, C. Poletto, J.J. Ramasco, N. Perra, M. Tizzoni, D. Paolotti, W. van den Broeck, A.J. Valleron, A. Vespignani, Modeling the critical care demand and antibiotics resources needed during the Fall 2009 wave of influenza A(H1N1) pandemic, *PLoS Curr.* 1 (December (DEC)) (2009) RRN1133.
- [28] O.M. Araz, J.W. Fowler, T.W. Lant, M. Jehn, A pandemic influenza simulation model for preparedness planning, in: Proc. Winter Simul. Conf., 2009, pp. 1986–1995.
- [29] R. Coker, Swine flu, *BMJ* 338 (2009) b1791.
- [30] V. Colizza, A. Vespignani, N. Perra, C. Poletto, B. Gonçalves, H. Hu, D. Balcan, D. Paolotti, W. Van den Broeck, M. Tizzoni, P. Bajardi, J.J. Ramasco, Estimate of Novel Influenza A/H1N1 cases in Mexico at the early stage of the pandemic with a spatially structured epidemic model, *PLoS Curr.* 1 (November (NOV)) (2009) RRN1129.
- [31] P.M.P. a Sloot, V. Coveney, G. Ertaylan, V. Müller, C. a Boucher, M. Bubak, HIV decision support: from molecule to man, *Philos. Trans. R. Soc. A* 367 (2009) 2691–2703.
- [32] D. Balcan, B. Gonçalves, H. Hu, J.J. Ramasco, V. Colizza, A. Vespignani, Modeling the spatial spread of infectious diseases: the GLobal Epidemic and Mobility computational model, *J. Comput. Sci.* 1 (August (3)) (2010) 132–145.
- [33] G. González-Parra, A.J. Arenas, D.F. Aranda, L. Segovia, Modeling the epidemic waves of AH1N1/09 influenza around the world, *Spat. Spatiotemporal Epidemiol.* 2 (December (4)) (2011) 219–226.
- [34] G. González-Parra, A.J. Arenas, B.M. Chen-Charpentier, A fractional order epidemic model for the simulation of outbreaks of influenza A(H1N1), *Math. Methods Appl. Sci.* 37 (October (15)) (2014) 2218–2226.
- [35] G. González-Parra, R.-J. Villanueva, J. Ruiz-Baragaño, J.-A. Moraño, Modelling influenza A(H1N1) 2009 epidemics using a random network in a distributed computing environment, *Acta Trop.* 143 (2015) 29–35.
- [36] W. Van den Broeck, C. Giannini, B. Gonçalves, M. Quaggiotto, V. Colizza, A. Vespignani, The GLEaMviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale, *BMC Infect. Dis.* 11 (January (1)) (2011).
- [37] X. Li, X. Li, Y.-Y. Jin, A Data-driven inference algorithm for epidemic pathways using surveillance reports in 2009 outbreak of influenza A (H1N1), in: 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), 2012, pp. 2840–2845.
- [38] Z. Cheng, X. Qiu, P. Zhang, R. Meng, *System Simulation and Scientific Computing*, vol. 326, CCIS, no. Part 1, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.
- [39] M. Tizzoni, P. Bajardi, C. Poletto, J.J. Ramasco, D. Balcan, B. Gonçalves, N. Perra, V. Colizza, A. Vespignani, Real-time numerical forecast of global epidemic spreading: case study of 2009 A/H1N1pdm, *BMC Med.* 10 (January) (2012).
- [40] N. Zarrabi, M. Prosperi, R.G. Belleman, M. Colafogli, A. De Luca, P.M.A. Sloot, Combining epidemiological and genetic networks signifies the importance of early treatment in HIV-1 transmission, *PLoS ONE* 7 (January (9)) (2012) e46156.
- [41] W. Duan, R. Quax, M. Lees, X. Qiu, P.M.A. Sloot, Topology dependent epidemic spreading velocity in weighted networks, *J. Stat. Mech. Theory Exp.* 2014 (December (12)) (2014) P12020.
- [42] S. Mei, Y. Zhu, X. Qiu, X. Zhou, Z. Zu, A.V. Boukhanovsky, P.M.A. Sloot, Individual decision making can drive epidemics: a fuzzy cognitive map study, *IEEE Trans. Fuzzy Syst.* 22 (April (2)) (2014) 264–273.
- [43] T.-Y. Tsai, K.-W. Chang, C.Y.-C. Chen, iScreen: world's first cloud-computing web server for virtual screening and de novo drug design based on TCM database@Taiwan, *J. Comput. Aided. Mol. Des.* 25 (June (6)) (2011) 525–531.
- [44] S. Zhou, R. Liao, J. Guan, When cloud computing meets bioinformatics: a review, *J. Bioinform. Comput. Biol.* 11 (October (50)) (2013) 1330002.
- [45] R. Lestari, M. Ahmad, B. Alisjahbana, T. Djatmiko, The lung diseases diagnosis software: influenza and tuberculosis case studies in the cloud computing environment, in: 2012 International Conference on Cloud Computing and Social Networking (ICCCSN), 2012, pp. 1–7.
- [46] H.P. Shanahan, A.M. Owen, A.P. Harrison, Bioinformatics on the cloud computing platform Azure, *PLOS ONE* 9 (January (7)) (2014) e102642.
- [47] L. Griebel, H.-U. Prokosch, F. Köpcke, D. Toddernoth, J. Christoph, I. Leb, I. Engel, M. Sedlmayr, A scoping review of cloud computing in healthcare, *BMC Med. Inform. Decis. Mak.* 15 (March (1)) (2015).
- [48] R. Sandhu, S.K. Sood, G. Kaur, An intelligent system for predicting and preventing MERS-CoV infection outbreak, *J. Supercomput.* (2015), <http://dx.doi.org/10.1007/s11227-015-14740>.
- [49] S.K. Sood, R. Sandhu, Matrix based proactive resource provisioning in mobile cloud environment, *Simul. Model. Pract. Theory* 50 (2014) 83–95.
- [50] 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.
- [51] E. Elsebakhi, F. Lee, E. Schendel, A. Haque, N. Kathireason, T. Pathare, N. Syed, R. Al-Ali, Large-scale machine learning based on functional networks for biomedical big data with high performance computing platforms, *J. Comput. Sci.* (September) (2015), <http://dx.doi.org/10.1016/j.jocs.2015.09.008>.
- [52] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA data mining software, *ACM SIGKDD Explor. Newsl.* 11 (1) (2009).
- [53] J. Scott, *Social Network Analysis*, vol. 19, SAGE Publications, 2012.
- [54] Amazon, AWS – Amazon EC2 Instance Types, 2014 [Online]. Available <http://aws.amazon.com/ec2/instance-types/> (accessed 19.07.15).
- [55] S. Tisue, U. Wilensky, NetLogo: design and implementation of a multi-agent modeling environment, *Proc. Agent* 2004 (2004) 7–9.
- [56] “UCI Machine Learning Repository: Adult Data Set.” [Online]. Available <https://archive.ics.uci.edu/ml/datasets/Adult> (accessed 04.08.15).



Rajinder Sandhu obtained his M.E. (Software Engineering) from Thapar University, Patiala in 2013 and B.Tech with Distinction from MMEC, Mullana in 2011. Presently he is pursuing his Ph.D. in cloud computing from Guru Nanak Dev University, Amritsar. His research areas are cloud computing, Virtualization and Software Engineering.



Harsuminder Kaur obtained her Bachelor of Technology from PTU, Jalandhar in 2013 and Master of Technology from Guru Nanak Dev University in 2015. Her area of interest are cloud computing and Big Data.



Dr. Sandeep K. Sood did his Phd in Computer Science & Engineering from IIT Roorkee, India. He is currently working as Associate Dean (A.A. & S.W), Head & Associate Professor, Computer Science & Engineering, G.N.D.U. Regional Campus, Gurdaspur. He has 16 years of teaching and 9 years of research experience. He has more than 50 research publication. His work is published and cited in highly reputed journals of Elsevier, Springer, Wiley and IEEE. His citation number according to Google Scholar is 427 with h-index equal to 11 and i10-index equal to 11. His research areas are Security, cloud computing and big data.